AJS

merican
ournal
f Sociology
olume 84 Number 4
nuary 1979

The AMERICAN JOURNAL OF SOCIOLOGY is published bimonthly in July, September, November, January, March, and May by The University of Chicago at The University of Chicago Press, 5801 Ellis Avenue, Chicago, Illinois 60637. Subscription rates, U.S.A.: institutions, 1 year $24.00, 2 years $43.20, 3 years $61.20; individuals, 1 year $18.00, 2 years $32.40, 3 years $45.90. Student subscription rate, U.S.A.: 1 year $14.40 (letter from professor must accompany subscription). ASA and BSA members, 1 year $14.40. All other countries add $2.00 for each year's subscription to cover postage. Single copy rates: institutions $5.00, individuals $4.00. Back issues available from 1962 (vol. 68). Subscriptions are payable in advance. Subscriptions will be entered to start with the first issue to be published after receipt of order. *Claims for missing numbers* should be made within the month following the regular month of publication. The publishers expect to supply missing numbers free only when losses have been sustained in transit and when the reserve stock will permit.

Second-class postage paid at Chicago, Illinois. PRINTED IN U.S.A.

 ISSN 0002-9602

AJS

American Journal of Sociology

Volume 84 No. 4 January 1979

CONTENTS

Commentary and Debate

Review Essays

Book Reviews

IN THIS ISSUE

OTIS DUDLEY DUNCAN is professor of sociology at the University of Arizona. He is coauthor, with Beverly Duncan, of the recently published monograph, *Sex Typing and Social Roles*. Their current research on social change and age stratification features a number of structural models for categorical data.

LEO A. GOODMAN is the Charles L. Hutchinson Distinguished Service Professor of Statistics and Sociology at the University of Chicago, and a research associate at the Population Research Center of the University. His book *Analyzing Qualitative/Categorical Data: Log-Linear Models and Latent-Structure Analysis* was published in 1978. A second book, *New Methods for the Analysis of Cross-Classification Tables: Goodman's Models*, was also published in 1978, translated into Hungarian by the Hungarian Academy of Sciences, Institute of Sociology.

NANCY BRANDON TUMA is assistant professor of sociology at Stanford University. Her main interests are stratification, economic structures, and the development of models and methods for studying change. She is currently engaged in research on labor mobility, marital stability, and change in social networks.

MICHAEL T. HANNAN is associate professor of sociology at Stanford University. His research interests include organizational ecology, social change, and welfare policy and the family. *National Development and the World System*, a collection of research with John W. Meyer and students, will be published soon. He is also completing (with Nancy Brandon Tuma) a monograph on models and methods for studying change.

LYLE P. GROENEVELD is a sociologist at the Center for the Study of Welfare Policy, SRI International. In addition to working on investigating the impact of the Seattle and Denver Income-Maintenance Experiments on marital stability and family functioning, he is directing a study of the relationship between family income and delinquency using data from the income-maintenance experiments.

WARREN HANDEL is associate professor of sociology at Southern Illinois University at Edwardsville. He is coauthor, with Robert Lauer, of *Social Psychology: The Theory and Application of Symbolic Interactionism*. He has eclectic interests in microsociology.

GARY KLECK is an instructor in the School of Criminology, Florida State University. His research interests are in violence, gun ownership, and variation in criminal sentencing by race.

TERRY S. WEINER is assistant professor and chairman of the Department of Sociology at Union College. He has published in the areas of political socialization and (with Eugene Schneller) minorities in the health professions. He has also been active in the undergraduate education section of the ASA and has published working papers on curriculum for the Project on Teaching Sociology. He is continuing his work on higher education

with particular emphasis on attitude change of minorities admitted to professional schools.

BRUCE K. ECKLAND is professor of sociology at the University of North Carolina at Chapel Hill. He will be spending the next year on a theoretical analysis of educational equity theory and the biological basis of individual and group differences in educational attainment.

CHARLES C. LEMERT is associate professor of sociology at Southern Illinois University at Carbondale. His research interests include theory and semiotics, and he has a book coming out entitled *Sociological Theory and the Twilight of Man.*

STEPHEN COLE is professor of sociology at the State University of New York at Stony Brook and is spending the year at the Center for Advanced Study in the Behavioral Sciences. He is currently completing a study for the National Academy of Sciences on the NSF peer review system and writing a series of articles analyzing the ways in which knowledge grows in the different sciences.

JOHN W. MEYER is professor of sociology at Stanford University. His research interests are studies in comparative sociology, sociology of education, and formal organizations.

KRZYSZTOF ZAGÓRSKI is a professor at the Institute of Philosophy and Sociology of the Polish Academy of Sciences. He has been involved in a number of empirical studies of Polish society. He is especially interested in social mobility and completed a large-scale study of mobility in Polish society while at the Central Statistical Office.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to nonsubstantive editing with the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his article without charge in any book of which he is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the journal has the non-exclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the journal in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, and references—double-spaced, allowing generous margins at top, bottom, and sides of page. Number all footnotes, including acknowledgment footnote.

2. Type each table on a separate page. Refer to each table in numerical order in the text. Prepare tables without vertical lines.

3. Clarify all symbols with words in the margin of the manuscript. Encircle these and other explanatory notes not intended for printing.

4. Draw figures on white paper with india ink. Original or glossy print of each figure will be required if manuscript is accepted.

5. Include an abstract of not more than 100 words summarizing the findings.

6. *Four copies* of the manuscript and abstract, on white paper, must be submitted. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 1130 East 59th Street, The University of Chicago, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose

a self-addressed stamped postcard, containing the manuscript title, to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article or research note submitted to the *AJS* must be accompanied by a check or money order for $10.00, payable to The University of Chicago Press in U.S. currency or its equivalent by postal or express money order or bank draft. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for student authors. Submissions from students should be accompanied by a faculty member's certification of their student status.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons.

Format of References: List all items alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "References." Examples follow:

Davis, K. 1963*a*. "Social Demography." Pp. 124–37 in *The Behavioral Sciences Today*, edited by Bernard Berelson. New York: Basic.

———. 1963*b*. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29 (October): 345–66.

Goode, W. J. 1967. "The Protection of the Inept." *American Sociological Review* 32 (February): 5–19.

Moore, Wilbert E., and Arnold S. Feldman. 1960. *Labor Commitment and Social Change in Developing Areas*. New York: Social Science Research Council.

Sanford, Nevitt, ed. 1962. *The American College*. New York: Wiley.

Weber, Max. (1921) 1968. "Society's Problems." Pp. 12–16 in *Economy and Society*, edited by Guenther Roth and Claus Wittich. New York: Bedminster.

AJS

American Journal of Sociology

Volume 84 No. 5 March 1979

CONTENTS

Review Essays

Book Reviews

IN THIS ISSUE

PRISCILLA P. CLARK is associate professor of French at the University of Illinois at Chicago Circle. She is the author of *The Battle of the Bourgeois* and is currently writing a book on literary institutions and literary culture in France.

LEO A. GOODMAN is the Charles L. Hutchinson Distinguished Service Professor of Statistics and Sociology at the University of Chicago, and a research associate at the Population Research Center of the University. His book *Analyzing Qualitative/Categorical Data: Log-Linear Models and Latent-Structure Analysis* was published in 1978. A second book, *New Methods for the Analysis of Cross-Classification Tables: Goodman's Models,* was also published in 1978, translated into Hungarian by the Hungarian Academy of Sciences, Institute of Sociology.

DAVID SNYDER is associate professor in the Department of Sociology, Indiana University. In addition to further analyses of world system structure, he is also currently engaged in research projects in the areas of collective action and violence and occupational sex segregation.

EDWARD L. KICK is currently an instructor in the Department of Sociology, University of Utah. Most recently he has been working on a longitudinal study of world-system structure and the influence of world-system properties on military interventions and national attitudes and experiences such as conflict and violence.

MICHAEL S. LEWIS-BECK is associate professor of political science at the University of Iowa. His areas of interest include political economy, public policy, and quantitative research methods. Currently he is examining the impact of economic conditions on presidential popularity.

DAVID P. PHILLIPS is associate professor of sociology at the University of California, San Diego. He is currently continuing his studies on suicide and disguised suicide, the impact of the mass media on violence, and a sociological theory of imitation.

WILLIAM A. RUSHING is chairman of the Department of Sociology and Anthropology, Vanderbilt University. His current research interests include the study of mental patient populations, with particular attention to the determinants and consequences of types of admission (voluntary vs. involuntary). He is also studying the psychological, social, and economic consequences of disability for individuals and families.

SUZANNE T. ORTEGA is a Ph.D. candidate in sociology at Vanderbilt University. Her major areas of interest are deviance and social psychology. She is currently studying the effects of physical disability and status change on the psychosocial integration of urban and rural adults.

Barry Wellman is associate professor of sociology at the University of Toronto and a member of the Centre for Urban and Community Studies. He is spending this year at the Netherlands Institute for Advanced Study. He is currently engaged in research combining the survey data discussed in this issue with longitudinal focused-interview information in order to study the structure, use, meaning, and change of urban social networks in Toronto.

John F. Myles is assistant professor of sociology at Carleton University, Ottawa. His current research focuses on the political economy of income maintenance programs for the aged and their role in reproducing and altering the structure of intracohort and intercohort inequality.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to nonsubstantive editing with the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his article without charge in any book of which he is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the journal has the non-exclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the journal in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, and references—double-spaced, allowing generous margins at top, bottom, and sides of page. Number all footnotes, including acknowledgment footnote.

2. Type each table on a separate page. Refer to each table in numerical order in the text. Prepare tables without vertical lines.

3. Clarify all symbols with words in the margin of the manuscript. Encircle these and other explanatory notes not intended for printing.

4. Draw figures on white paper with india ink. Original or glossy print of each figure will be required if manuscript is accepted.

5. Include an abstract of not more than 100 words summarizing the findings.

6. *Four copies* of the manuscript and abstract, on white paper, must be submitted. Manuscripts should be addressed to the Editor of the American Journal of Sociology, 1130 East 59th Street, The University of Chicago, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose

a self-addressed stamped postcard, containing the manuscript title, to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article or research note submitted to the *AJS* must be accompanied by a check or money order for $10.00, payable to The University of Chicago Press in U.S. currency or its equivalent by postal or express money order or bank draft. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for student authors. Submissions from students should be accompanied by a faculty member's certification of their student status.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons.

Format of References: List all items alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "References." Examples follow:

Davis, K. 1963*a*. "Social Demography." Pp. 124–37 in *The Behavioral Sciences Today*, edited by Bernard Berelson. New York: Basic.

———. 1963*b*. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29: 345–66.

Goode, W. J. 1967. "The Protection of the Inept." *American Sociological Review* 32: 5–19.

Moore, Wilbert E., and Arnold S. Feldman. 1960. *Labor Commitment and Social Change in Developing Areas*. New York: Social Science Research Council.

Sanford, Nevitt, ed. 1962. *The American College*. New York: Wiley.

Weber, Max. (1921) 1968. "Society's Problems." Pp. 12–16 in *Economy and Society*, edited by Guenther Roth and Claus Wittich. New York: Bedminster.

A CALL FOR PAPERS

for a Supplement to the *American Journal of Sociology*
on Marxist inspirations in contemporary social research

Editors: Michael Burawoy and Theda Skocpol

This supplement is to contain primarily research-based articles that will demonstrate the fascination and fruitfulness of Marxist ideas in contemporary social research. The editors are looking for concrete studies that are conceived and executed within some Marxist tradition *or* address a problem distinctively raised by Marxist writers *or* test or develop Marxist ideas about a given aspect of social life. Papers may address topics in any of a wide range of areas, such as the class structures of advanced societies, the family, urban or community studies, the labor process, politics and the state, gender or race and ethnic relations, world capitalism or imperialism, transitions between modes of production, social history, and ideology or culture. Submissions will be welcome from both established and junior scholars in sociology and in history, economics, political science, and anthropology.

Procedures: Papers must be submitted for preliminary consideration no later than February 1, 1980. They may be no longer than 45 typewritten, double-spaced pages and should be prepared according to the "Information for Contributors" in the *American Journal of Sociology*. *Three copies* of each submission should be mailed *either* to Michael Burawoy, Department of Sociology, University of California, Berkeley, California 94720, *or* to Theda Skocpol, Department of Sociology, William James Hall, Harvard University, Cambridge, Massachusetts 02138.

Papers will be reviewed by both editors, and those that appear potentially suitable will be sent anonymously for advisory review to one or two qualified referees. Papers may be accepted subject to substantive or stylistic revisions.

How Destination Depends on Origin in the Occupational Mobility Table[1]

Otis Dudley Duncan
University of Arizona

A reanalysis of the expanded version of the 1949 British cross-classification table of respondent's occupation by father's occupation features two models that emphasize the "achievement" rather than the "mobility" point of view hitherto taken in contingency-table analysis of such data. The new models are attractive in terms of mathematical parsimony, statistical goodness of fit, conceptual clarity, and suitability for application in comparative work. The favorable results are attributed partly to the fairly substantial degree of disaggregation of occupation categories. The models capture some of the flavor of the regression approach to mobility in that they show in a straightforward manner how the destination distribution is shifted for each shift of one or more steps on the scale of occupational origins.

In his insightful review of structural models for mobility tables, Hauser (1978) alludes to the tension between the regression approach to the analysis of occupational achievement and the contingency-table approach to the analysis of occupational mobility. Pullum (1975) makes an even sharper distinction between analyses using the general linear model for quantitative variables and those employing log-linear models for the frequencies in cross-classification tables. While there are several ways in which the two approaches have diverged, the salient one here is the distinction between models which emphasize social origins and other background characteristics as factors in *achievement* and those which focus on the comparison of origin and destination statuses in order to characterize patterns of occupational *inheritance and mobility*. It is characteristic of regression models (e.g., Duncan and Hodge 1963; Blau and Duncan 1967) that occupational status is treated as a scaled variable, so that dependence of destination on origin is represented, say, by a simple or partial regression coefficient. (The assumption of linearity, of course, can be relaxed.) Many of the contingency-table analyses (e.g., Goodman 1965; White 1970; Boudon 1975; Pullum 1975), on the other hand, take as given a set of occupational classes and inquire whether the class of destination is the same as or different from the class of origin, so that special attention is given to "occupational inheri-

[1] The assistance of Douglas M Sloane and the suggestions of Leo A Goodman, Robert M. Hauser, and Hal H. Winsborough are gratefully acknowledged.

tance," or "status immobility," or the latent classes of "stayers" versus "movers."

Goodman (1972) has shown how to introduce the idea of "distance" between occupational classes into the contingency-table approach. Models involving this notion presuppose that occupational classes are scaled with respect to status. But the models treated by Goodman (1972) retain the emphasis on inheritance and mobility, inasmuch as the relevant distance is the one traversed in moving from occupational origin *i* to occupational destination *j*. In most of Goodman's models the diagonal of the mobility table is given special treatment. And the parameters of various models are described as taking account of "downward" versus "upward mobility," or as representing effects of "status inertia" or "status barriers." Similarly, Hauser (1978), while not emphasizing occupational inheritance to the same degree as the other writers cited, claims that contingency-table methods are useful in telling "us where in the social structure opportunities for movement or barriers to movement are greater or less, and in so doing provide clues about stratification processes which are not less important, if different in kind, from those uncovered by multivariate causal models" (p. 921).

The purpose of this inquiry is to consider the possibility of combining features of the two approaches. Ideally, one would like both to retain the elegance of the regression description of the origin-by-destination table, which expresses the dependence of achievement on background in terms of one or a few regression coefficients, and to maintain the fidelity to the detailed pattern of cell frequencies that is characteristic of the contingency-table approach. To demonstrate the feasibility of doing just that, let us reexamine the classic table of Glass (1954), in the modified form published by Miller (1960, p. 71), which provides a subdivision of Glass's occupational class V, distinguishing nonmanual from manual workers in that class. The expanded classification is as follows:

I. Professional and high administrative
II. Managerial and executive
III. Inspectional, supervisory, and other nonmanual (high grade)
IV. Inspectional, supervisory, and other nonmanual (lower grade)
Va. Routine grades of nonmanual
Vb. Skilled manual
VI. Semiskilled manual
VII. Unskilled manual

The top section of table 1 reproduces the raw data.

For expository purposes I discuss five models altogether. They are listed in table 2, but the order of discussion will differ from the order of listing in the table. Because this paper does not attempt to be mathematically rigorous or to achieve a high level of mathematical generality, it will suffice to

characterize each model by the constraint on the fitted counts (expected frequencies) which it incorporates.[2] As for notation, we are studying an R

TABLE 1

BRITISH MOBILITY TABLE, 1949, WITH FITTED COUNTS FOR TWO MODELS AND DIAGONAL RATIOS FOR FIVE MODELS

FATHER'S OCCUPATION	SON'S OCCUPATION I	II	III	IV	Va	Vb	VI	VII
	Observed Counts							
I ($i=1$)	50	19	26	8	7	11	6	2
II (2)	16	40	34	18	11	20	8	3
III (3)	12	35	65	66	35	88	23	21
IV (4)	11	20	58	110	40	183	64	32
Va (5)	2	8	12	23	25	46	28	12
Vb (6)	12	28	102	162	90	554	230	177
VI (7)	0	6	19	40	21	158	143	71
VII (8)	0	3	14	32	15	126	91	106
	Fitted Counts, Model (4)							
I	9.4	14.0	22.2	18.0	5.5	14.4	3.6	1.4
II	9.0	15.3	27.8	25.8	9.0	27.3	7.8	3.4
III	15.6	30.4	63.4	67.6	27.0	93.6	30.7	15.1
IV	13.4	30.0	71.7	87.6	40.0	159.3	59.8	33.8
Va	2.2	5.6	15.2	21.3	11.1	50.8	21.9	14.1
Vb	10.4	30.7	96.3	154.5	92.6	484.8	239.2	177.2
VI	1.7	5.6	20.1	36.9	25.4	152.2	86.1	73.1
VII	0.7	2.8	11.8	24.8	19.6	134.3	87.0	84.6
	Fitted Counts, Model (5)							
I	10.5	15.8	23.3	17.7	5.2	13.0	3.1	1.1
II	13.2	20.9	32.8	26.4	8.2	21.8	5.5	2.1
III	15.3	31.5	64.2	67.0	27.1	93.6	30.6	14.9
IV	9.6	24.5	61.3	78.8	39.2	166.4	67.0	40.1
Va	2.1	5.8	15.4	21.1	11.2	50.8	21.8	14.0
Vb	10.6	32.8	99.7	155.4	93.9	484.2	236.7	172.1
VI	1.6	5.6	19.8	36.0	25.4	152.7	87.1	73.9
VII	0.8	3.2	12.5	25.4	20.0	134.0	85.2	80.6

	Diagonal Ratios: Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
I	13.2	1.6	34.7	5.3	4.8
II	5.9	1.6	8.6	2.6	1.9
III	2.0	1.1	2.2	1.0	1.0
IV	1.6	1.2	1.7	1.3	1.4
Va	2.3	2.0	2.3	2.2	2.2
Vb	1.2	1.1	1.3	1.1	1.1
VI	1.8	1.3	2.2	1.7	1.6
VII	2.3	1.2	2.9	1.3	1.3

SOURCE.—Miller 1960, p. 71.

[2] See the companion paper, Goodman (1979), for a more general mathematical treatment of the models proposed here as well as some models considered earlier (Goodman 1972).

TABLE 2

TEST STATISTICS FOR FIVE MODELS

Model	df	X^2
(1) Independence ..	$(R-1)^2=49$	954 5
(2) Row effects	$(R-1)(R-2)=42$	130 1
(3) Quasi independence, diagonal omitted	$(R-1)^2-R=41$	446 8
(4) Uniform association, diagonal omitted* ..	$R^2-3R=40$	58 4
(5) Row effects, diagonal omitted*	$R^2-4R+2=34$	45 9

*Model is nontrivial only if $R \geq 4$

$\times R$ table with observed counts f_{ij} ($i = 1, \ldots, R$; $j = 1, \ldots, R$) and fitted counts F_{ij} under a specified model. Diagonal counts are f_{ii} and F_{ii}, respectively.

The standard baseline model of independence (or perfect mobility) fits the row and column marginals and requires that

$$F_{ij}F_{i'j'}/F_{i'j}F_{ij'} = 1.0\,. \tag{1}$$

No assumption is made concerning the ordering of rows or columns.

The model of quasi independence (or quasi-perfect mobility), proposed by Goodman (1965), may be defined in various ways, according to the selection of cells put aside for separate treatment. Here we consider only the case in which all the diagonal cells are omitted from the calculation. The model fits the row and column marginals (obtained after the diagonal counts are set equal to zero) and requires that

$$\begin{gathered} F_{ij}F_{i'j'}/F_{i'j}F_{ij'} = 1.0 \\ i \neq j\,, \quad i' \neq j\,, \quad i \neq j'\,, \quad i' \neq j'\,. \end{gathered} \tag{3}$$

The only assumption about ordering is the trivial one that rows and columns are in the same (arbitrary) order, so that the diagonal is uniquely defined.

A simple but far-reaching modification of the quasi-independence model is made by specifying order and introducing a single additional parameter, which must be estimated from the data. As in model (3), row and column marginals, exclusive of diagonal cell counts, are fitted; but there is a new constraint:

$$\begin{gathered} F_{ij}F_{i+1,j+1}/F_{i+1,j}F_{i,j+1} = b \\ i \neq j\,, \quad i \neq j+1\,, \quad i \neq j-1\,, \end{gathered}$$

from which it may be deduced that

$$\begin{gathered} F_{ij}F_{i'j'}/F_{i'j}F_{ij'} = b^{(i-i')(j-j')} \\ i \neq j\,, \quad i' \neq j\,, \quad i \neq j'\,, \quad i' \neq j'\,. \end{gathered} \tag{4}$$

This model requires that the rows and columns be ordered in the "correct"

way, since the model is defined only after the ordering is specified. It will be termed the model of "uniform association."[3] Its parameter, b, has an obvious conceptual analogy to the regression coefficient, since it governs the shift of the destination distribution that occurs for each "step" of the ordered origin classification. But b is not, of course, a regression coefficient in the usual sense of the word.

Simon (1974) proposed a model for the singly ordered contingency table that he labels "Formulation A." The columns of the table are ordered but the rows need not be. That is, the model ignores any information about the ordering of rows. This model will here be termed the "row-effects model," since it requires a set of parameters, b_i, one for each row, of which only $R - 1$ are independent, however. The constraint on the fitted frequencies is given by

$$F_{ij}F_{i'j'}/F_{i'j}F_{ij'} = (b_i/b_{i'})^{(j'-j)} . \quad (2)$$

Finally, we consider a model which is affiliated to quasi independence in much the same way as the row-effects model is affiliated to independence. That is, row effects are specified only for the nondiagonal cells, the diagonal cells being set aside:

$$\text{Model (2) with } i \neq j , \quad i' \neq j , \quad i \neq j' , \quad i' \neq j' . \quad (5)$$

Let us turn to methods of fitting the models[4] and testing their goodness of fit. Since the treatment of models (1) and (3) in the literature is quite full, there is no need to dwell on the details for them. It suffices to note that the values of the likelihood-ratio χ^2 statistic (X^2) in table 2 are large for these models, so that neither model would be seriously considered for other than didactic uses.

The method of iterative scaling (used, for example, in fitting model [3]) can be modified, as Simon (1974) observes, so as to fit not only the row sums $\Sigma_j f_{ij}$ and the column sums $\Sigma_i f_{ij}$ but also the weighted sum of column numbers $\Sigma_j j f_{ij}$. The constraint cited above as a definition of model (2) is satisfied when Simon's algorithm for it is used. With his procedure we obtain $X^2 = 130.1$. This compares favorably with the fit achieved on the hypothesis of quasi independence even though model (2) gives no special attention to the diagonal.

Model (4) can be fitted using a variant of the procedure of iterative

[3] H. H. Winsborough reminds me that uniform association (over the whole table) is a strong form of the property of isotropy found in contingency tables which are constructed by forming class intervals for a bivariate normal distribution (Yule 1906) The occupational mobility table, however, is more like the table "produced by taking an isotropic distribution and then increasing the frequencies in the diagonal compartment by a sufficient percentage" (Yule and Kendall 1950, p 59).

[4] Other methods of fitting these models are presented in the companion paper (Goodman 1979). The discussion here is retained for its possible heuristic value.

scaling. One specifies a trial value of the parameter b and then obtains fitted counts and X^2 making use of the "starting table" feature of the computer program Everyman's Contingency Table Analyzer (ECTA) (Fay and Goodman 1973). A search then ensues for alternative values of $\hat{b}$ which will produce lower values of X^2. The search terminates when one finds the value that virtually minimizes X^2. For $\hat{b} = 1.146$ we obtain $X^2 = 58.45$; for $\hat{b} = 1.148$, $X^2 = 58.46$; and for $\hat{b} = 1.140$, $X^2 = 58.94$. A slight improvement in the estimate we accepted ($\hat{b} = 1.146$) is perhaps possible.

An iterative procedure was devised for fitting model (5). The initial step was the fitting of model (2). Fitted diagonal counts from that model were substituted for the observed diagonal counts in the input to the next round of fitting, using the Simon algorithm. This process was repeated until there was an acceptable approximation to convergence, that is, when the fitted diagonal counts at stage s were practically the same as they had been at stage $s - 1$. Since the steps were carried out as a sequence of batch problems, we were somewhat liberal as to the convergence criterion. Some slight improvement in the F_{ij} for model (5), as shown in table 1, would perhaps be possible with further calculation. It should be noted that the effect of this estimation procedure is to render the diagonal cells irrelevant to the F_{ij} ultimately obtained for the remaining cells and to the value of X^2.

Calculation of degrees of freedom for the several models is straightforward once one has enumerated the independent parameters estimated or constraints fitted, including, of course, those attributable to the fitting of row and column marginals. Formulas are shown in table 2. Hierarchical relationships (Goodman 1970) among the models are illustrated in figure 1. We may treat the difference in X^2 between two hierarchical models as a χ^2 statistic, with degrees of freedom equal to the difference in degrees of freedom of the two models. This provides a test of whether the additional parameters fitted in the more complicated model yield a significant improvement in fit compared with the simpler model.

We find that model (2), like models (1) and (3), does not afford a satisfactory fit. For model (4), the uniform-association model with diagonal omitted, we find that $.025 < P < .05$, so that the model is close to being acceptable on a conventional criterion. For model (5), the row-effects model with diagonal omitted, we actually secure an acceptable fit on that criterion, $.05 < P < .1$. There are, of course, alternative measures of closeness of fit that have been proposed, but there is no need to invoke them here. A statistical curiosity turns up in the hierarchical comparison of the last two models, for $X^2(M_4) - X^2(M_5) = 12.5$, with df $= 40 - 34 = 6$, $.05 < P < .1$, so that model (5) does not actually afford a significant improvement (at the .05 level) over model (4), although model (5) itself has an acceptable fit (at the .05 level) and model (4) does not. In this comparison of

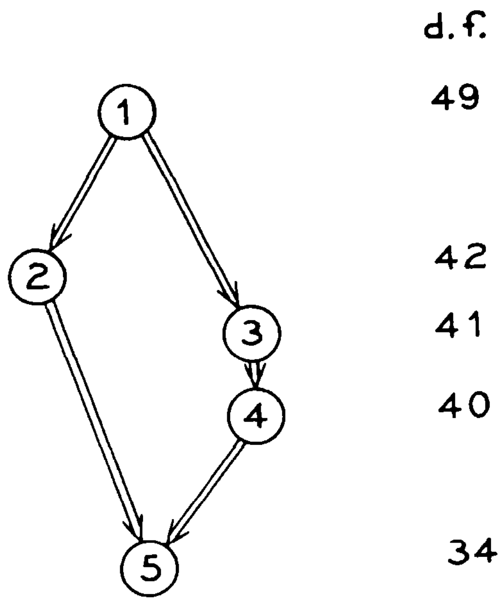


FIG. 1.—Hierarchical relationships among models. The double-shafted arrow means "implies"; implication is transitive. Goodman (1973) suggests this kind of diagram.

the two models we are in a cliff-hanger situation which, in principle, could only be resolved with more data.[5]

The reason for this situation is apparent from figure 2. The row effects under model (5) are very nearly a log-linear function of row number, while model (4) produces row effects that are an exact log-linear function of row number. Indeed, it was the unexpectedly close approximation to linearity for model (5) that suggested model (4). In model (5), it should be noted, there is no constraint which ensures that estimated values of the row-effect parameters will exhibit a monotonic, let alone an almost linear, pattern. Moreover, the order of the occupational classes was not selected ex post facto to produce this pattern. It was determined by previous investigators on the basis of considerations about occupational status that have no necessary relationship to the properties of our model. In short, figure 2 records an empirical discovery, albeit one that may not be surprising in the light of Blau and Duncan's demonstration that "socioeconomic status is a fundamental dimension of the social distance between occupational groups" (1967, p. 74).

To complete the description of results obtained with these models, we turn to the matter of the diagonal. Under models (1) and (2), the diagonal

[5] A reader objects that this predicament "is in fact quite common and should not be referred to as [a curiosity]." Perhaps the curiosity is that the statisticians have so little to say about this "quite common" occurrence.

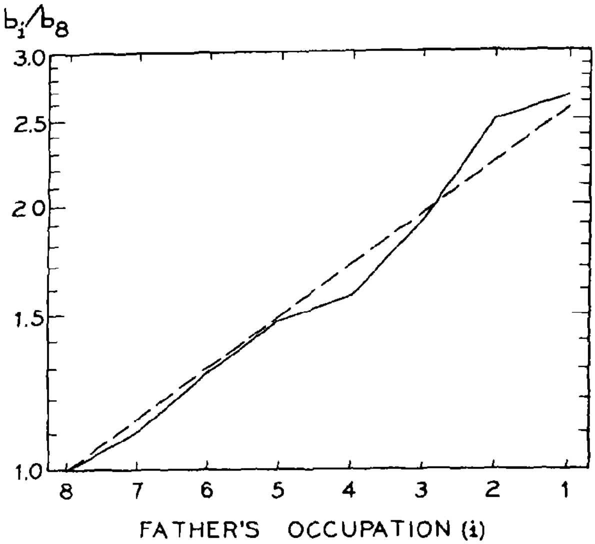


FIG. 2.—Parameter estimates for model (4) (dashed line) and model (5) (solid line).

cells are included in the set for which fitted counts are obtained, so that they contribute—and, typically, contribute heavily—to X^2. Moreover, the model, in "trying to fit" the diagonal, is led to adopt poorly fitting counts for the off-diagonal cells. In the remaining three models this problem is avoided, at the price of complicating the model, insofar as a separate treatment of the diagonal cells is required. One way (not the only way) to handle them is to extend the model—originally defined only for off-diagonal cells—to the diagonal, while holding fixed the estimates of parameter values based on the off-diagonal cells alone (Goodman 1969, 1972). As already noted, the procedure for fitting model (5) accomplishes this in the process of fitting. For model (3), once the off-diagonal F_{ij} have been found, we may compute implied or imputed diagonal counts F^*_{ii} $(i = 1, \ldots, R)$ by the formula $F^*_{ii} = F_{i'i}F_{ij'}/F_{i'j'}$ $(i \neq j', i' \neq j')$, which is obtained by letting $i = j$ in the constraint for model (3). Analogously, for model (4) we use the formula $F^*_{ii} = b^{(i-i')(i-j')}F_{i'i}F_{ij'}/F_{i'j'}$.

The diagonal ratios in the bottom section of table 1, accordingly, refer to f_{ii}/F_{ii} as far as models (1) and (2) are concerned and to f_{ii}/F^*_{ii} in the case of model (3), (4), or (5). For reasons that have been given in the literature (e.g., Goodman 1969; Hauser 1978), it is not a good idea to place substantive interpretations on diagonal ratios pertaining to a model that does not fit acceptably well. We pass over the attractively low ratios for model (2), therefore.

The patterns of diagonal ratios for models (4) and (5) are similar. Both

show the relatively high value for occupational class I that has often been noted for these data in the past. Neither set, however, has the especially high ratio at the bottom of the occupational scale that one might have expected from model (1). Moreover, neither model uncovers evidence of "status disinheritance." That is, we find no case of $f_{ii}/F_{ii} < 1.0$ or $f_{ii}/F^*_{ii} < 1.0$. It is interesting that under the model of quasi independence—model (3) in this paper—Goodman (1969) observed status disinheritance when the British table was collapsed to three rows and columns. But with the same data in the less highly aggregated form of a 5×5 or 7×7 table, Goodman (1972) did not find diagonal ratios under 1.0 when the model of quasi independence was under consideration. Some of the other models he studied did, however, manifest this effect (Goodman 1972, tables IX and X). The comparison of the results obtained with different levels of aggregation, therefore, supports the argument for relatively detailed occupational classes, inasmuch as "status disinheritance" may be a difficult process to explain.

This treatment of the British table is, obviously, only an illustrative one. But, as Hauser remarks, "These data . . . have been so thoroughly analyzed . . . that they have become a standard set of observations against which each new model or method is calibrated" (1978, p. 932). Results obtained here do suggest that an important limitation of previous analyses has been the use of too highly aggregated versions of table 1. Certainly the substantive plausibility of model (4) or (5) is enhanced by using a finer grading of occupational status. In any event, though, what we can learn from a single mobility table is decidedly limited. Hope's comment is well taken: "Where the techniques come into their own, even when they are applied to hazy categories, is in *comparative* analysis" (1976, p. 730). Model (4) is especially attractive for comparative research, for the comparison there will take the very sharp form of testing the hypothesis that the model's one parameter for row-by-column interaction is the same in two or more populations, or that it varies over time or over populations according to some function that the researcher might aspire to specify.

It is not quite so clear how one uses the diagonal ratios in comparative investigations, although one can certainly make informal and possibly illuminating contrasts of the magnitudes and patterns of variation over occupational classes, in much the same way that one can study residuals from the fitted counts of the models. Perhaps the next development will be some kind of model that conserves the attractive features of model (4) but handles the diagonal in a less ad hoc fashion. Further disaggregation of the occupational categories should help, for, as disaggregation proceeds, the diagonal declines in importance (Duncan 1966; Blau and Duncan 1967, pp. 111, 142, 173).

Models (4) and (5) give similar answers to the question, How does destination depend on origin in the occupational mobility table? The dependence takes two forms. One form is represented by the parameter or set of parameters that in a simple and clearly specified manner shifts the destination distribution upward or downward as the origin is shifted up or down. Second, there is a residual tendency, not fully accounted for by either model, for the destination class to be the same as the class of origin. As is shown in the companion paper (Goodman 1979), models (4) and (5), which stress the dependence of achievement on origin, can be specified mathematically in a form similar to that taken by other models which emphasize mobility, or the difference between destination and origin statuses. Thus, the choice between the two points of view may be made within the more general class of multiplicative structural models for the cross-classification table. It need not commit the analyst to a simultaneous choice between linear and log-linear models.

REFERENCES

Blau, P M., and O. D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Boudon, R. 1975. "A Model for the Analysis of Mobility Tables." Pp. 511–27 in *Quantitative Sociology*, edited by H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, and V. Capecchi. New York: Academic Press

Duncan, O. D 1966 "Methodological Issues in the Analysis of Social Mobility " Pp. 51–97 in *Social Structure and Mobility in Economic Development*, edited by N. J. Smelser and S M. Lipset. Chicago: Aldine.

Duncan, O D., and R. W. Hodge. 1963. "Education and Occupational Mobility." *American Journal of Sociology* 68 (May): 629–44.

Fay, R., and L. A Goodman 1973. "ECTA Program Description for Users." Unpublished document available from Leo A. Goodman, University of Chicago

Glass, D V., ed. 1954 *Social Mobility in Britain* London· Routledge & Kegan Paul

Goodman, L. A 1965. "On the Statistical Analysis of Mobility Tables." *American Journal of Sociology* 70 (March): 564–85.

———. 1969. "How to Ransack Social Mobility Tables and Other Kinds of Cross-Classification Tables " *American Journal of Sociology* 75 (July): 1–39.

——— 1970. "The Multivariate Analysis of Qualitative Data: Interactions among Multiple Classifications." *Journal of the American Statistical Association* 65 (March)· 226–56.

——— 1972. "Some Multiplicative Models for the Analysis of Cross-classified Data." Pp. 649–96 in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, edited by L. Le Cam, J. Neyman, and E. L. Scott. Berkeley: University of California Press

———. 1973 "Guided and Unguided Methods for the Selection of Models for a Set of T Multidimensional Contingency Tables." *Journal of the American Statistical Association* 68 (March). 165–75.

———. 1979. "Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-Classification Tables." *American Journal of Sociology* 84 (January)· 804–19

Hauser, R. M. 1978. "A Structural Model of the Mobility Table." *Social Forces* 56 (March): 919–53.

Hope, K 1976. Review of Pullum (1975). *American Journal of Sociology* 82 (November): 726–31.
Miller, S. M. 1960. "Comparative Social Mobility." *Current Sociology* 9 (1): 1–89.
Pullum, T. 1975. *Measuring Occupational Inheritance*. New York: Elsevier.
Simon, G. 1974. "Alternative Analyses for the Singly-ordered Contingency Table" *Journal of the American Statistical Association* 69 (December): 971–76.
White, H. C. 1970. "Stayers and Movers." *American Journal of Sociology* 76 (September): 307–24.
Yule, G. U. 1906. "On a Property Which Holds Good for All Groupings of a Normal Distribution of Frequency for Two Variables, with Applications to the Study of Contingency Tables for the Inheritance of Unmeasured Qualities" *Proceedings of the Royal Society*, ser A 77: 324–36.
Yule, G. U., and M. G. Kendall. 1950. *An Introduction to the Theory of Statistics*. 14th ed. New York: Hafner

Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-Classification Tables[1]

Leo A. Goodman
University of Chicago

This note shows how some specific models proposed by Duncan (1979), Goodman (1972), Haberman (1974), and Simon (1974), which turn out to be useful for the analysis of occupational mobility tables and other kinds of cross-classification tables, are related to models obtained by applying the general formulations of multiplicative models in Goodman (1972), Darroch and Ratcliff (1972), and Haberman (1974). It also shows how each model can be described in terms of its expected frequencies, or equivalently in terms of its pattern of odds-ratios, or equivalently in terms of its pattern of odds. In so doing, it provides (*a*) further insight into the meaning of some of the specific models, (*b*) greater flexibility in the possible interpretations of a wide range of potentially useful specific models included within the general class of multiplicative models developed for the analysis of cross-classification tables, and (*c*) improved methods for estimating the expected frequencies under some of the specific models.

In the analysis of occupational mobility tables describing, say, the cross-classification of son's occupational category and father's occupational category, we will usually find that the baseline model of perfect mobility (i.e., independence between son's occupational category and father's occupational category) will not serve as an adequate description of the observed data. A more adequate description can usually be obtained by modifying the baseline model in ways that can take account of the possible existence of, say, "occupational inheritance and/or disinheritance," "occupational immobility and/or persistence," "overall upward or downward occupational mobility," "occupational mobility inertia," "occupational mobility barriers," and other related phenomena, using models of the kind introduced in, for example, Goodman (1972). With these models, which can take account of the possible existence of various relevant substantive phenomena, the statistical methods presented in Goodman (1972) can be used (*a*) to determine whether a satisfactory description of the observed data

[1] This research was supported in part by research contract no. NSF SOC76-80389 from the Division of the Social Sciences of the National Science Foundation For helpful comments, the author is indebted to O. D. Duncan, C. C. Clogg, and R. M. Hauser.

can be obtained by taking account of these phenomena, (*b*) to estimate the magnitude of the effects of these phenomena, (*c*) to determine which of these effects are statistically significant and which are not, and (*d*) to test various hypotheses about the effects of these phenomena.

These phenomena will affect various aspects of the observed data in various ways. They will affect (*a*) the pattern of the observed frequencies in the cells of the cross-classification table (i.e., the occupational mobility table); (*b*) the observed pattern of association between the column classification (i.e., son's occupational category) and the row classification (i.e., father's occupational category) as measured, say, by the pattern of the odds-ratios (i.e., cross-product ratios) calculated from 2×2 subtables of the cross-classification table; (*c*) the pattern of the observed odds that an observation will fall in column category j' rather than in category j'', (i.e., that the son will be in occupational category j' rather than in category j''), given that it was in row category i (i.e., given that the father was in occupational category i); and (*d*) the pattern of the observed odds that an observation would fall in row category i' rather than in category i'' (i.e., that the father would be in occupational category i' rather than in category i''), given that it is in column category j (i.e., given that the son is in occupational category j). Each of the models introduced in Goodman (1972) could be described in terms of its expected frequencies, or equivalently in terms of its pattern of odds-ratios, or equivalently in terms of its pattern of odds. The present note provides general formulae that show how a model expressed in terms of its expected frequencies can be reexpressed in terms of its pattern of odds-ratios or in terms of its pattern of odds; how a model expressed in terms of its pattern of odds-ratios can be reexpressed in terms of its expected frequencies or in terms of its pattern of odds; and how a model expressed in terms of its pattern of odds can be reexpressed in terms of its expected frequencies or in terms of its pattern of odds-ratios. (See formulae [19a]-[19c], [27a]-[27b], [28a]-[28b], [32], [33a]-[33b], [34a]-[34b] below.) These general formulae can be applied to a specific model originally described using one form of expression (say, in terms of its expected frequencies) in order to obtain descriptions (or interpretations) of the same model using other forms of expression. They can also be applied, in comparing specific models originally described using different forms of expression, in order to determine whether or not the models are equivalent, and if not in which respects they differ from each other. Application of these general formulae also can serve as a heuristic device to assist the research worker in his search for models that are particularly well suited to a given substantive area and in his examination of the possible interpretations and implications of these models.

The above remarks referring to the analysis of occupational mobility tables can be applied more generally to the analysis of other kinds of cross-classification tables as well. In the occupational mobility table there is a

one-to-one correspondence between the classes of the row classification and the classes of the column classification, and the classes of the row (and column) classification can usually be ordered (from high to low). Although some of the specific models which will be considered here are particularly well suited to square contingency tables of this kind (in which there is a one-to-one correspondence and in which the classes of the row [and column] classification can usually be ordered), the general results presented in the present note and in Goodman (1972) can be applied also to rectangular tables (as well as to square tables), where there may or may not be some kind of correspondence between the classes of the row and column classifications and where the classes of the row (and/or column) classification may or may not be ordered. These results can also be directly extended to the analysis and comparison of two (or more) $R \times C$ cross-classification tables and to the analysis of multidimensional cross-classification tables.

For expository purposes, the general results will be presented near the end of this note. The note will consider first some simple specific models of the kind introduced in Goodman (1972). Although these models could be described for more general cross-classification tables, for simplicity of exposition our attention will be focused on, say, a 4×4 table with an observed count f_{ij} ($i = 1, 2, 3, 4; j = 1, 2, 3, 4$) in the ith row and jth column of the table and a corresponding expected count F_{ij} under a specified model. The observed and expected counts are displayed in tables 1 and 2. Under the standard baseline model of independence (or perfect mobility), the F_{ij} are of the simple multiplicative form

$$F_{ij} = \alpha_i \beta_j \,, \tag{1}$$

where the α_i and β_j are positive constants pertaining to the ith row and jth column, respectively (see, e.g., Goodman 1965). The expected counts under condition (1) are displayed in table 3.

As noted earlier herein, the expected counts under the baseline model of independence will usually not serve as an adequate description of an observed table of counts, and when they do not we can consider a wide range of modifications of the baseline model (see, e.g., Goodman 1969, 1972). These modifications introduce various multiplicative constants to take account of the various inadequacies of the baseline model as a description of the observed table of counts. It may even be possible sometimes to obtain an adequate description with the introduction of a single multiplicative constant, say, γ, to take account of the inadequacy of the baseline model. For example, if the inadequacy pertains primarily to the counts in the cells on the main diagonal, it may be possible to obtain an adequate description by introducing a single "inflation factor" γ, applied to the cells on the main

diagonal, to take account of this inadequacy.[2] With the introduction of the multiplicative constant γ, formula (1) for the expected count F_{ij} is replaced by

$$F_{ij} = \begin{cases} \alpha_i\beta_j , & \text{for } i \neq j \\ \alpha_i\beta_j\gamma , & \text{for } i = j . \end{cases} \qquad (2)$$

The expected counts under condition (2) are displayed in table 4. This model is a simple example of a model included within the general class of

Table 1

f_{11}	f_{12}	f_{13}	f_{14}
f_{21}	f_{22}	f_{23}	f_{24}
f_{31}	f_{32}	f_{33}	f_{34}
f_{41}	f_{42}	f_{43}	f_{44}

Table 2

F_{11}	F_{12}	F_{13}	F_{14}
F_{21}	F_{22}	F_{23}	F_{24}
F_{31}	F_{32}	F_{33}	F_{34}
F_{41}	F_{42}	F_{43}	F_{44}

Table 3

$\alpha_1\beta_1$	$\alpha_1\beta_2$	$\alpha_1\beta_3$	$\alpha_1\beta_4$
$\alpha_2\beta_1$	$\alpha_2\beta_2$	$\alpha_2\beta_3$	$\alpha_2\beta_4$
$\alpha_3\beta_1$	$\alpha_3\beta_2$	$\alpha_3\beta_3$	$\alpha_3\beta_4$
$\alpha_4\beta_1$	$\alpha_4\beta_2$	$\alpha_4\beta_3$	$\alpha_4\beta_4$

Table 4

$\alpha_1\beta_1\gamma$	$\alpha_1\beta_2$	$\alpha_1\beta_3$	$\alpha_1\beta_4$
$\alpha_2\beta_1$	$\alpha_2\beta_2\gamma$	$\alpha_2\beta_3$	$\alpha_2\beta_4$
$\alpha_3\beta_1$	$\alpha_3\beta_2$	$\alpha_3\beta_3\gamma$	$\alpha_3\beta_4$
$\alpha_4\beta_1$	$\alpha_4\beta_2$	$\alpha_4\beta_3$	$\alpha_4\beta_4\gamma$

Table 5

γ	1	1	1
1	γ	1	1
1	1	γ	1
1	1	1	γ

Table 6

1	δ	δ	δ
δ	1	δ	δ
δ	δ	1	δ
δ	δ	δ	1

[2] In the substantive context of social mobility analysis, if there is, in a certain specific sense, "status inheritance" in each of the cells on the main diagonal, then the multiplicative constant γ is an "inflation factor"; whereas if there is "status disinheritance" in each of the cells on the main diagonal, then the multiplicative constant γ is a "deflation factor." When the amount of status inheritance and/or status disinheritance differs in the different cells on the main diagonal, a different multiplicative constant can be applied to each cell on the main diagonal (say, γ_i pertaining to the ith cell on the main diagonal), and the usual model of quasi independence (or quasi-perfect mobility) is obtained (see, e.g., Goodman 1965, 1972). The multiplicative constants γ_i correspond to the "new index of immobility" in Goodman (1969, 1972).

multiplicative models introduced in Goodman (1972),[3] and the various statistical techniques presented in that article can be applied here.

Both models (1) and (2) include the multiplicative constants α_i and β_j pertaining to the ith row and jth column, respectively (for $i = 1, 2, 3, 4$; $j = 1, 2, 3, 4$); and all other models considered herein will also include these constants. Since these constants will be included in all of the models, it will be convenient for expository purposes to focus attention explicitly on the other multiplicative constants (if any) in the models, with the constants α_i and β_j included implicitly. Thus, instead of displaying the expected counts under model (2) as in table 4, for convenience we shall use table 5 for the display pertaining to this model.[4]

The introduction of a single "inflation factor" γ applied to the cells on the main diagonal, as in model (2), would actually be equivalent to the introduction of a single "deflation factor," say δ, applied to the cells that are not on the main diagonal. Thus, the expected counts under model (2) could also be expressed in the following equivalent form:[5]

$$F_{ij} = \begin{cases} \alpha_i\beta_j\,, & \text{for } i = j \\ \alpha_i\beta_j\delta\,, & \text{for } i \neq j\,. \end{cases} \tag{3}$$

Table 6 provides the display pertaining to this model.[6]

Model (3) is a special case of the usual "diagonals-parameter model" (applied to the full cross-classification table) where the same multiplicative

[3] Model (2) is included within the general case described in section 3 of Goodman (1972). This model can also be viewed as a special kind of "diagonals-parameter model" applied to the full cross-classification table (see sections 2.5 and 2.8 in the 1972 article) where the diagonals-parameter pertains only to the cells on the main diagonal. This model is also equivalent to a special case of the usual diagonals-parameter model (for the full cross-classification table) where all the diagonals-parameters, except for the parameter pertaining to the main diagonal, are set equal to one (or set equal to each other).

[4] Tabular displays of this general kind, corresponding to the various models considered herein and other models as well (see, e.g., tables 5–13), can serve as a heuristic device to assist the research worker in his search for models that are particularly well suited to a given substantive area. This device and others introduced later in the present note (e.g., the description of specific models using three different but equivalent forms of expression), can provide the research worker with a "machine" for building models that take account of relevant substantive phenomena. A wide variety of models can be examined in this way.

[5] The methods presented in Goodman (1972) for showing that two particular multiplicative models are equivalent can be applied directly here to show that models (2) and (3) are equivalent. Models (2) and (3) describe the "uniform inheritance model" and/or "uniform disinheritance model."

[6] The multiplicative constant δ in table 6 (i.e., in model [3]) is equal to $1/\gamma$, where γ is the multiplicative constant in table 5 (i.e., in model [2]). Table 6 can be obtained from table 5 by dividing each entry in table 5 by γ. (The constants pertaining to the rows and columns im model [3] are also related to but different from the corresponding constants in model [2]. More specifically, the α_i and β_j in model [3] are actually equal to $\alpha_i\gamma\nu$ and β_j/ν, respectively, for the α_i, β_j, and γ in model [2], with ν denoting any positive constant of proportionality.)

constant δ pertains to all cells that are not on the main diagonal.[7] In other words, in model (3) the diagonals-parameters are set equal to each other for all cells (i, j) where $|i - j| \neq 0$. When the diagonals-parameters in the usual diagonals-parameter model differ in the different diagonals, and the difference depends upon how far the diagonal is from the main diagonal, the diagonals-parameter pertaining to cell (i, j) will depend upon $|i - j|$. Table 7 provides the display pertaining to this situation. In this case, the

Table 7

1	δ_1	δ_2	δ_3
δ_1	1	δ_1	δ_2
δ_2	δ_1	1	δ_1
δ_3	δ_2	δ_1	1

Table 8

1	δ	δ^2	δ^3
δ	1	δ	δ^2
δ^2	δ	1	δ
δ^3	δ^2	δ	1

Table 9

1	δ_1	δ_1^2	δ_1^3
1	δ_2	δ_2^2	δ_2^3
1	δ_3	δ_3^2	δ_3^3
1	δ_4	δ_4^2	δ_4^3

Table 10

1	δ	δ^2	δ^3
1	δ^2	δ^4	δ^6
1	δ^3	δ^6	δ^9
1	δ^4	δ^8	δ^{12}

Table 11

1	1	1	1
1	δ	δ^2	δ^3
1	δ^2	δ^4	δ^6
1	δ^3	δ^6	δ^9

Table 12

1	δ	δ^2	δ^3
δ	δ^3	δ^5	δ^7
δ^2	δ^5	δ^8	δ^{11}
δ^3	δ^7	δ^{11}	δ^{15}

Table 13

1	1	1	1
1	ζ	ζ	ζ
1	ζ	ζ^2	ζ^2
1	ζ	ζ^2	ζ^3

[7] For related comments, see n. 3 above.

expected count F_{ij} can be expressed as follows:

$$F_{ij} = \alpha_i \beta_j \delta_k , \quad \text{for } k = |i - j| \tag{4}$$

where $\delta_0 = 1$. As already noted for the models considered earlier herein, model (4) is also a simple example of a model included within the general class of multiplicative models introduced in Goodman (1972), and the various statistical techniques presented in that article can be applied to the model.[8]

Comparing tables 6 and 7, we see that model (3) is a special case of model (4) in which the following condition is satisfied:[9]

$$\delta_k = \delta , \quad \text{for } k = 1, 2, 3 . \tag{5}$$

Consider next the model obtained from model (4) when the δ_k satisfy the following condition (instead of condition [5]):

$$\delta_k = \delta^k , \quad \text{for } k = 1, 2, 3 . \tag{6}$$

In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^k , \quad \text{for } k = |i - j| . \tag{7}$$

Table 8 provides the display for this model.[10] Condition (6) and the corresponding model (7) were mentioned in Goodman (1972), and the model is equivalent, in most respects, to the "fixed distance model" introduced by Haberman (1974).[11]

[8] In the context of social mobility analysis, the diagonals-parameter δ_k measures "status inertia" pertaining to a move from status category i to status category $i + k$ or $i - k$ (for $i = 1, 2, 3, 4$), i.e., a move of k status categories. The magnitude δ_k of "status inertia" (for $k = 1, 2, 3$) can be estimated using model (4) or more general models that include diagonals-parameters and other relevant parameters (see Goodman 1972), in the case where model (4) (or a more general model) is congruent with the observed data. The subscript k in model (4) pertains to the kth subset of the cells in the cross-classification table, when the cells are partitioned into subsets so that the cells included in the same subset are on the same diagonal. Formula (4) was also applied by Hauser (1978) with the cells partitioned into subsets so that the cells included in the same subset have expected frequencies that are at the same density level. As with model (4), the model applied in Hauser (1978) is an example of the kind of model included in the general class of multiplicative models introduced in Goodman (1972).

[9] In the context of social mobility analysis, condition (5) states that the magnitude δ_k of "status inertia" pertaining to a move of k status categories is equal for $k = 1, 2, 3$. To test whether the "status inertia" magnitudes are equal, we can compare the goodness-of-fit of model (4) (or a more general model that includes diagonals-parameters and other relevant parameters) and model (3) (or the corresponding more general model with condition [5] satisfied), in the case where the former model is congruent with the observed data.

[10] Model (7) is related to, but different from, the "uniform association model" in Duncan (1979). If the absolute difference $|i - j|$ in the formula for model (7) is replaced by the product ij, as in the formula for model (14) presented later herein, we then obtain a model that is equivalent to the "uniform association model."

[11] Model (7) is a special case of the "crossings-parameter model," introduced in Goodman (1972), in which all the "crossings-parameters" are equal in magnitude. In the context

Note how the multiplicative constant in the jth cell of the first row of table 8 is raised to higher powers as j increases (for $j = 1, 2, 3, 4$). Consider next the model obtained when this kind of pattern (observed in the first row of table 8) is applied to each row of the table, with a different multiplicative constant applied to each row. Table 9 provides the display for the model thus obtained. In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta_i^{j-1} . \tag{8}$$

Model (8) is equivalent to the model that was presented in Simon (1974) and applied in Duncan (1979) to occupational mobility data.[12]

Consider next the model obtained from model (8) when the δ_i satisfy the following condition:

$$\delta_i = \delta^i , \quad \text{for } i = 1, 2, 3, 4 . \tag{9}$$

Table 10 provides the display for the model thus obtained. In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^{i(j-1)} . \tag{10}$$

Model (10) is also equivalent to the models corresponding to the displays in tables 11 and 12, and the corresponding expected count F_{ij} for these models can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^{(i-1)(j-1)} \tag{11}$$

and

$$F_{ij} = \alpha_i \beta_j \delta^{ij-1} , \tag{12}$$

respectively.[13] Indeed, models (10), (11), and (12) are all equivalent to the

---

of social mobility analysis, the "crossings-parameters" measure "status barriers" (or "distances") between adjacent status categories. The magnitudes of the "status barriers" (or "distances") can be estimated using the "crossings-parameter model" as in Goodman (1972), or the related "variable distance model" in Haberman (1974), or more general models that include crossings-parameters and other relevant parameters (see, e.g., Goodman 1972). To test whether "status barriers" (or "distances") between adjacent categories are equal in magnitude, we can compare the goodness-of-fit of the "crossings-parameter model" (or a more general model) and model (7) (or the corresponding more general model with all the crossings-parameters set equal), in the case where the former model is congruent with the observed data.

[12] See "formulation A" in Simon (1974) and the "row-effects model" in Duncan (1979). Further comments on model (8) are included later in the present note and in Goodman (1978).

[13] The methods presented in Goodman (1972) for showing that two particular models are equivalent can be applied directly here to show that models (10), (11), and (12) are equivalent. Table 11 can be obtained from table 10 by dividing each entry in the jth column of table 10 by δ^{j-1}, and table 12 can be obtained by multiplying each entry in the ith row by δ^{i-1}. (The constants pertaining to the rows and columns in model [10] are related to but different from the corresponding constants in model [11], which are in turn related to but different from the corresponding constants in model [12]. For example, the α_i and β_j in model [10] are actually equal to $\alpha_i \nu$ and $\beta_j / \delta^{j-1} \nu$, respectively, for the α_i, β_j, and δ in model [11], with ν denoting any positive constant of proportionality.)

expected count F_{ij} can be expressed as follows:

$$F_{ij} = \alpha_i \beta_j \delta_k , \quad \text{for } k = |i - j| \tag{4}$$

where $\delta_0 = 1$. As already noted for the models considered earlier herein, model (4) is also a simple example of a model included within the general class of multiplicative models introduced in Goodman (1972), and the various statistical techniques presented in that article can be applied to the model.[8]

Comparing tables 6 and 7, we see that model (3) is a special case of model (4) in which the following condition is satisfied:[9]

$$\delta_k = \delta , \quad \text{for } k = 1, 2, 3 . \tag{5}$$

Consider next the model obtained from model (4) when the δ_k satisfy the following condition (instead of condition [5]):

$$\delta_k = \delta^k , \quad \text{for } k = 1, 2, 3 . \tag{6}$$

In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^k , \quad \text{for } k = |i - j| . \tag{7}$$

Table 8 provides the display for this model.[10] Condition (6) and the corresponding model (7) were mentioned in Goodman (1972), and the model is equivalent, in most respects, to the "fixed distance model" introduced by Haberman (1974).[11]

[8] In the context of social mobility analysis, the diagonals-parameter δ_k measures "status inertia" pertaining to a move from status category i to status category $i + k$ or $i - k$ (for $i = 1, 2, 3, 4$), i.e , a move of k status categories. The magnitude δ_k of "status inertia" (for $k = 1, 2, 3$) can be estimated using model (4) or more general models that include diagonals-parameters and other relevant parameters (see Goodman 1972), in the case where model (4) (or a more general model) is congruent with the observed data. The subscript k in model (4) pertains to the kth subset of the cells in the cross-classification table, when the cells are partitioned into subsets so that the cells included in the same subset are on the same diagonal. Formula (4) was also applied by Hauser (1978) with the cells partitioned into subsets so that the cells included in the same subset have expected frequencies that are at the same density level. As with model (4), the model applied in Hauser (1978) is an example of the kind of model included in the general class of multiplicative models introduced in Goodman (1972).

[9] In the context of social mobility analysis, condition (5) states that the magnitude δ_k of "status inertia" pertaining to a move of k status categories is equal for $k = 1, 2, 3$ To test whether the "status inertia" magnitudes are equal, we can compare the goodness-of-fit of model (4) (or a more general model that includes diagonals-parameters and other relevant parameters) and model (3) (or the corresponding more general model with condition [5] satisfied), in the case where the former model is congruent with the observed data

[10] Model (7) is related to, but different from, the "uniform association model" in Duncan (1979). If the absolute difference $|i - j|$ in the formula for model (7) is replaced by the product ij, as in the formula for model (14) presented later herein, we then obtain a model that is equivalent to the "uniform association model."

[11] Model (7) is a special case of the "crossings-parameter model," introduced in Goodman (1972), in which all the "crossings-parameters" are equal in magnitude. In the context

Note how the multiplicative constant in the jth cell of the first row of table 8 is raised to higher powers as j increases (for $j = 1, 2, 3, 4$). Consider next the model obtained when this kind of pattern (observed in the first row of table 8) is applied to each row of the table, with a different multiplicative constant applied to each row. Table 9 provides the display for the model thus obtained. In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta_i^{j-1} . \tag{8}$$

Model (8) is equivalent to the model that was presented in Simon (1974) and applied in Duncan (1979) to occupational mobility data.[12]

Consider next the model obtained from model (8) when the δ_i satisfy the following condition:

$$\delta_i = \delta^i , \quad \text{for } i = 1, 2, 3, 4 . \tag{9}$$

Table 10 provides the display for the model thus obtained. In this case, the expected count F_{ij} can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^{i(j-1)} . \tag{10}$$

Model (10) is also equivalent to the models corresponding to the displays in tables 11 and 12, and the corresponding expected count F_{ij} for these models can be expressed as

$$F_{ij} = \alpha_i \beta_j \delta^{(i-1)(j-1)} \tag{11}$$

and

$$F_{ij} = \alpha_i \beta_j \delta^{ij-1} , \tag{12}$$

respectively.[13] Indeed, models (10), (11), and (12) are all equivalent to the

---

of social mobility analysis, the "crossings-parameters" measure "status barriers" (or "distances") between adjacent status categories The magnitudes of the "status barriers" (or "distances") can be estimated using the "crossings-parameter model" as in Goodman (1972), or the related "variable distance model" in Haberman (1974), or more general models that include crossings-parameters and other relevant parameters (see, e g , Goodman 1972). To test whether "status barriers" (or "distances") between adjacent categories are equal in magnitude, we can compare the goodness-of-fit of the "crossings-parameter model" (or a more general model) and model (7) (or the corresponding more general model with all the crossings-parameters set equal), in the case where the former model is congruent with the observed data.

[12] See "formulation A" in Simon (1974) and the "row-effects model" in Duncan (1979) Further comments on model (8) are included later in the present note and in Goodman (1978).

[13] The methods presented in Goodman (1972) for showing that two particular models are equivalent can be applied directly here to show that models (10), (11), and (12) are equivalent. Table 11 can be obtained from table 10 by dividing each entry in the jth column of table 10 by δ^{j-1}, and table 12 can be obtained by multiplying each entry in the ith row by δ^{i-1}. (The constants pertaining to the rows and columns in model [10] are related to but different from the corresponding constants in model [11], which are in turn related to but different from the corresponding constants in model [12]. For example, the α_i and β_j in model [10] are actually equal to $\alpha_i \nu$ and $\beta_j / \delta^{j-1} \nu$, respectively, for the α_i, β_j, and δ in model [11], with ν denoting any positive constant of proportionality)

models obtained when the expected count can be expressed as[14]

$$F_{ij} = \alpha_i \beta_j \delta^{(i-1)j} , \tag{13}$$

or as

$$F_{ij} = \alpha_i \beta_j \delta^{ij} , \tag{14}$$

or as

$$F_{ij} = \alpha_i \beta_j \delta^{(i-2.5)(j-2.5)} . \tag{15}$$

Models (10)–(15) are all equivalent to the "uniform association model" in Duncan (1979).[15]

Having noted above the equivalence of models (10)–(15), we comment next on the fact that the other models presented earlier herein can also be expressed in various equivalent forms. For example, models (2) and (3) (namely, the "uniform inheritance model" and "uniform disinheritance model") are equivalent, as we noted earlier (see tables 5 and 6). Similarly, the "fixed distance model" as described by model (7) (see table 8) is equivalent to the model corresponding to the display in table 13, and the corresponding expected count F_{ij} can be expressed as follows:[16]

$$F_{ij} = \begin{cases} \alpha_i \beta_j \zeta^{i-1} , & \text{for } i \leq j \\ \alpha_i \beta_j \zeta^{j-1} , & \text{for } i > j . \end{cases} \tag{16}$$

We may also note here that model (8) (see table 9) is equivalent to the models

[14] The remarks in n. 13 can be applied also to models (13), (14), and (15). It may also be worth noting that the equivalence of the models is not affected by replacing 2.5, which appears twice in the exponent in the formula for model (15), by any other specified numbers (e.g., by 1 in model [11] and by 0 in model [14]); and the specified numbers used in the two places in the exponent in the formula for model (15) can differ from each other (e.g., the 2.5 appearing first in the exponent in the formula for model [15] is replaced by 0 in model [10] and by 1 in model [13]; and the 2.5 appearing second in the exponent is replaced by 1 in model [10] and by 0 in model [13]). In some contexts, it may be convenient to use as the specified numbers $(R + 1)/2$ and $(C + 1)/2$, where R and C denote the number of rows and columns, respectively, in the cross-classification table.

[15] The formulation of this model presented in the present note will lead us later herein to an iterative estimation procedure that is different from the one used by Duncan (1979) to study this model (On this topic, see also Goodman 1978.) Furthermore, because of the relationship between the "uniform association model" and the "fixed distance model" (model [7]) described in n. 10, the general iterative estimation procedure referred to later herein can be applied also to the "fixed distance model" (and other related models) as well as to the "uniform association model."

[16] The remarks in n. 13 can be applied also to models (7) and (16). The multiplicative constant ζ in table 13 (i.e., in model [16]) is equal to $1/\delta^2$, where δ is the multiplicative constant in table 8 (i.e., in model [7]). Table 13 can be obtained from table 8 by dividing each entry in the jth column of table 8 by δ^{j-1} and each entry in the ith row by δ^{i-1}. Models (7) and (16) might appear, at first sight, to be quite different from each other, but their equivalence will become more evident when we discuss the corresponding odds-ratios later herein. (Although table 13 might appear, at first sight, to resemble table 11 more closely than table 8, the model corresponding to table 13 is not equivalent to the model corresponding to table 11, but it is equivalent to the model corresponding to table 8.)

obtained when the expected count can be expressed as[17]

$$F_{ij} = \alpha_i \beta_j \delta_i^{\,j} , \tag{17}$$

or as

$$F_{ij} = \alpha_i \beta_j \delta_i^{\,j-2.5} . \tag{18}$$

Each of the models considered herein can be described in terms of its corresponding pattern of odds-ratios (i.e., cross-product ratios). In order to demonstrate this, we now shall consider the quantity Θ_{ij} defined as follows:

$$\Theta_{ij} = (F_{ij}F_{i+1,j+1})/(F_{i,j+1}F_{i+1,j}) \tag{19a}$$

$$= (F_{ij}/F_{i,j+1})/(F_{i+1,j}/F_{i+1,j+1}) \tag{19b}$$

$$= (F_{ij}/F_{i+1,j})/(F_{i,j+1}/F_{i+1,j+1}) . \tag{19c}$$

The above formulae define the odds-ratio (i.e., the cross-product ratio) Θ_{ij} in the following equivalent terms: (*a*) It is the cross-product ratio obtained from the expected counts (i.e., the F's) in the 2×2 subtable formed from the cells (i, j), $(i, j + 1)$, $(i + 1, j)$, $(i + 1, j + 1)$; (*b*) it is the ratio of the odds $F_{ij}/F_{i,j+1}$ that an observation will fall in column j rather than in column $j + 1$, given that it was in row i, and the odds $F_{i+1,j}/F_{i+1,j+1}$ that an observation will fall in column j rather than in column $j + 1$, given that it was in row $i + 1$; and (*c*) it is the ratio of the odds $F_{i,j}/F_{i+1,j}$ that an observation would fall in row i rather than in row $i + 1$, given that it is in column j, and the odds $F_{i,j+1}/F_{i+1,j+1}$ that an observation would fall in row i rather than in row $i + 1$, given that it is in column $j + 1$.[18]

By applying formula (19a) to model (1) (see table 3), we of course obtain

$$\Theta_{ij} = 1 , \quad \text{for all } i, j . \tag{20}$$

Similarly, the odds-ratio for, say, model (7) (see table 8) can be expressed as follows:[19]

$$\Theta_{ij} = \begin{cases} 1/\delta^2 , & \text{for } i = j \\ 1 , & \text{for } i \neq i . \end{cases} \tag{21}$$

[17] The remarks in n 13 can be applied to models (8), (17), and (18) It may also be worth noting that the equivalence of the models is not affected by replacing 2.5, which appears in the exponent in the formula for model (18), by any other specified number (e.g., by 1 in model [8], and by 0 in model [17]). In some contexts, it may be convenient to use as the specified number $(C + 1)/2$, where C denotes the number of columns in the cross-classification table.

[18] For simplicity, only those odds-ratios that pertain to the cells in two successive rows (say, rows i and $i + 1$) and two successive columns (say, cols. j and $j + 1$) are considered here. All other odds-ratios pertaining to the cross-classification table can be calculated directly from this particular set of odds-ratios (see, e.g., Goodman 1969, 1978) This particular set can also be used to determine the formula for the expected counts F_{ij} under the model and the formulae for the corresponding odds (see formulae [32] and [33a]–[33b] below). Since this particular set can be used in determining the F_{ij} under the model, it can also be used to determine the corresponding tabular display (see, e.g., tables 5–13).

[19] The corresponding odds-ratio for model (16) (see table 13) is equal to ζ when $i = j$, and it is equal to 1 when $i \neq j$. This result helps to explain why models (7) and (16) are equivalent, with $\zeta = 1/\delta^2$.

More generally, the odds-ratio for model (4) (see table 7) can be expressed as

$$\Theta_{ij} = \begin{cases} 1/\delta_1^2 , & \text{for } i = j \\ \delta_k^2/(\delta_{k-1}\delta_{k+1}) , & \text{for } i \neq j , \end{cases} \tag{22}$$

with $k = |i - j|$ and $\delta_0 = 1$. (This odds-ratio for model [4] can, of course, also be written as

$$\Theta_{ij} = v_k , \quad \text{for } k = |i - j| \tag{23}$$

with $v_0 = 1/\delta_1^2$ and $v_k = \delta_k^2/(\delta_{k-1}\delta_{k+1})$, for $k = 1, 2, \ldots$.) In particular, for models (2) and (3) (see, e.g., table 6), we obtain the following:[20]

$$\Theta_{ij} = \begin{cases} 1/\delta^2 , & \text{for } k = 0 \\ \delta , & \text{for } k = 1 \\ 1 , & \text{for } k = 2, 3, \ldots . \end{cases} \tag{24}$$

Consider next model (8) (see table 9). For this model, the odds-ratio can be expressed as follows:[21]

$$\Theta_{ij} = \delta_{i+1}/\delta_i , \quad \text{for all } i, j . \tag{25}$$

And for models (10)–(15) (see tables 10–12), we obtain[22]

$$\Theta_{ij} = \delta , \quad \text{for all } i, j . \tag{26}$$

Formulae (20)–(26) above describe the pattern of odds-ratios for each of the models considered herein. Each of these models can also be described in terms of its corresponding pattern of odds (rather than odds-ratios). In order to demonstrate this, consider now the quantities $\Omega_{ij}^{A\bar{B}}$ and $\Omega_{ji}^{B\bar{A}}$ defined as follows:

$$\Omega_{ij}^{A\bar{B}} = F_{ij}/F_{i,j+1} , \tag{27a}$$

$$\Omega_{ji}^{B\bar{A}} = F_{ij}/F_{i+1,j} . \tag{27b}$$

The above formulae define the odds $\Omega_{ij}^{A\bar{B}}$ and the odds $\Omega_{ji}^{B\bar{A}}$ in the following terms: (*a*) $\Omega_{ij}^{A\bar{B}}$ is the odds $F_{ij}/F_{i,j+1}$ that an observation will fall in column j rather than in column $j + 1$, given that it was in row i; and (*b*) $\Omega_{ji}^{B\bar{A}}$ is the

[20] Formula (24) can be calculated directly from table 6, and it also follows as a special case of formula (22), since $\delta_k = \delta$ (for $k = 1, 2, \ldots$) in models (2) and (3) as noted in formula (5). Similarly, formula (21) can be calculated directly from table 8 (or table 13), and it also follows as a special case of formula (22), since $\delta_k = \delta^k$ in model (7) as noted in formula (6)

[21] It was noted earlier that model (8) was equivalent to a model presented in Simon (1974) and Duncan (1979) In the latter article, the model is defined in terms of the odds-ratios; in the former, in terms of the odds. In the present note, the model is defined in terms of the expected counts F_{ij} since this definition provides further insight into the meaning of the model, and it will facilitate discussion below of iterative estimation procedures for this model and other related models. For further comments on this topic, see Goodman (1978).

[22] Formula (26) can be calculated directly from table 10 (or tables 11 and 12), and it also follows as a special case of formula (25), since $\delta_i = \delta^i$ in models (10)–(15) as noted in formula (9).

odds $F_{ij}/F_{i+1,j}$ that an observation would fall in row i rather than in row $i+1$, given that it is in column j.[23] From (19b)-(19c) and (27a)-(27b), we obtain

$$\Theta_{ij} = \Omega_{ij}^{A\bar{B}}/\Omega_{i+1,j}^{A\bar{B}} \quad (28a)$$

$$= \Omega_{ji}^{B\bar{A}}/\Omega_{j+1,i}^{B\bar{A}} . \quad (28b)$$

By applying (27a)-(27b) to model 1 (see table 3), we of course obtain

$$\Omega_{ij}^{A\bar{B}} = \psi_j^B , \quad (29a)$$

$$\Omega_{ji}^{B\bar{A}} = \psi_i^A , \quad (29b)$$

with $\psi_i^A = \alpha_i/\alpha_{i+1}$ and $\psi_j^B = \beta_j/\beta_{j+1}$. Similarly, the odds for, say, models (10)-(15) (see, e.g., table 12) can be expressed as follows:[24]

$$\Omega_{ij}^{A\bar{B}} = \psi_j^B \xi^i , \quad (30a)$$

$$\Omega_{ji}^{B\bar{A}} = \psi_i^A \xi^j , \quad (30b)$$

where $\xi = 1/\delta$. More generally, the odds for model (8) (see table 9) can be expressed as[25]

$$\Omega_{ij}^{A\bar{B}} = \psi_j^B \xi_i , \quad (31a)$$

$$\Omega_{ji}^{B\bar{A}} = \psi_i^A \tau_i^j , \quad (31b)$$

where $\xi_i = 1/\delta_i$ and $\tau_i = \delta_i/\delta_{i+1}$. The other models presented earlier can also be expressed in similar terms, but these details will be left as an exercise for the interested reader.

We have now seen how each model can be described in terms of the pattern of odds-ratios (i.e., the Θ_{ij}), or equivalently in terms of the pattern of odds (i.e., the $\Omega_{ij}^{A\bar{B}}$ or the $\Omega_{ji}^{B\bar{A}}$), or equivalently in terms of the pattern of expected frequencies (i.e., the F_{ij}). For example, the "uniform association model" states that the odds-ratios Θ_{ij} are equal for all i and j (see formula [26]). And this model is equivalent to the model obtained by stating that the odds $\Omega_{ij}^{A\bar{B}}$ increase (or decrease) exponentially as row i increases (see

[23] For simplicity, only those odds $\Omega_{ij}^{A\bar{B}}$ that pertain to the cells in two successive columns (say, cols j and $j+1$) in a given row (say, row i), and those odds $\Omega_{ji}^{B\bar{A}}$ that pertain to the cells in two successive rows (say, rows i and $i+1$) in a given column (say, col j), are considered here All other odds that pertain to the cells in any two columns (say, cols. j' and j'') in row i can be calculated directly from the $\Omega_{ij}^{A\bar{B}}$ (as the product of the $\Omega_{ij}^{A\bar{B}}$'s for $j = j', j'+1, \ldots, j''-1$, when $j' < j''$); and all other odds that pertain to the cells in any two rows (say, rows i' and i'') in col. j can be calculated directly from the $\Omega_{ji}^{B\bar{A}}$ (as the product of the $\Omega_{ji}^{B\bar{A}}$'s for $i = i', i'+1, \ldots, i''-1$, when $i' < i''$).

[24] Formulae (30a) and (30b) remain correct even if the exponents i and j are replaced by i minus any specified number and j minus any specified number, respectively. The relationship between ψ_i^A and α_i/α_{i+1}, and between ψ_j^B and β_j/β_{j+1}, will depend upon which exponents are used in formulae (30a) and (30b), and upon which of the models (10)-(15) is used.

[25] The remarks in n. 24 can be applied also to the exponent j in formula (31b) and the corresponding ψ_i^A. Formulae similar to (31a) and (31b) were used in Simon (1974) to define a model that is equivalent to model (8), as noted in n. 21.

formula [30a]); and the rate of increase (or decrease) is unaffected by the value of j, but the intercept of the corresponding exponential curve may be affected by the value of j.[26] And this in turn is equivalent to the model obtained by stating that the odds $\Omega_{ji}^{B\bar{A}}$ increase (or decrease) exponentially as column j increases (see formula [30b]); and the rate of increase (or decrease) is unaffected by the value of i, but the intercept of the corresponding exponential curve may be affected by the value of i.[27] And this in turn is equivalent to the model obtained by stating that the expected frequency F_{ij} is the product of a row factor α_i, a column factor β_j, and an interaction factor that increases (or decreases) exponentially as the product ij increases (see formula [14]).[28]

The equivalent descriptions in the preceding paragraph pertain to models (10)–(15). A similar set of equivalent descriptions can be obtained for, say, model (8) using formulae (25), (31a), and (31b). Similar results can also be obtained for each of the other models presented earlier herein.

For any model described in terms of the expected frequencies F_{ij}, formula (19) can be used to calculate the corresponding odds-ratios Θ_{ij}, and formulae (27a) and (27b) can be used to calculate the corresponding odds $\Omega_{ij}^{A\bar{B}}$ and $\Omega_{ji}^{B\bar{A}}$, respectively. For any model described in terms of the odds-ratios, we can calculate the corresponding expected frequencies from[29]

$$F_{ij} = \alpha_i\beta_j\prod_{s=1}^{i-1}\prod_{t=1}^{j-1}\Theta_{st}\,; \tag{32}$$

and we can calculate the corresponding odds from

$$\Omega_{ij}^{A\bar{B}} = (\beta_j/\beta_{j+1}) \Big/ \prod_{s=1}^{i-1}\Theta_{sj}\,, \tag{33a}$$

[26] Expressed in terms of logarithms, formula (30a) can be used to describe the following linear relationship: The logarithm of the odds $\Omega_{ij}^{A\bar{B}}$ increases (or decreases) linearly as row i increases; and the slope of the increase (or decrease) is unaffected by the value of j, but the intercept of the corresponding straight line may be affected by the value of j.

[27] Formula (30b) can be used to describe a linear relationship for the logarithm of the odds $\Omega_{ji}^{B\bar{A}}$ in the same manner used in n. 26 to describe a linear relationship for the logarithm of the odds $\Omega_{ij}^{A\bar{B}}$.

[28] Expressed in terms of logarithms, formula (14) can be used to describe the following relationship: The logarithm of the expected frequency F_{ij} is the sum of a row factor ($\log \alpha_i$), a column factor ($\log \beta_j$), and an interaction factor that increases (or decreases) linearly as the product ij increases.

[29] The products

$$\prod_{s=1}^{i-1}\Theta_{st} \quad \text{and} \quad \prod_{t=1}^{j-1}\Theta_{st}$$

in formula (32) are clearly defined for the integer i (where $i = 2, 3, \ldots, R$) and the integer j (where $j = 2, 3, \ldots, C$), respectively; and when $i = 1$ or $j = 1$ the corresponding symbols can be defined as being equal to one.

$$\Omega_{ji}^{B\bar{A}} = (\alpha_i/\alpha_{i+1}) \Big/ \prod_{l=1}^{j-1} \Theta_{il} , \tag{33b}$$

where the α_i and β_j are the parameters used in formula (32). For any model described in terms of the odds, we can calculate the corresponding expected frequencies from[30]

$$F_{ij} = \alpha_i \Big/ \prod_{l=1}^{j-1} \Omega_{il}^{A\bar{B}} \tag{34a}$$

or

$$F_{ij} = \beta_j \Big/ \prod_{s=1}^{i-1} \Omega_{js}^{B\bar{A}} , \tag{34b}$$

and we can calculate the corresponding odds-ratios from formulae (28a)-(28b). It may also be worthwhile to note that the parameters α_i and β_j in formula (32) can be calculated in terms of the corresponding odds:

$$\alpha_i = \alpha_1 \Big/ \prod_{s=1}^{i-1} \Omega_{1s}^{B\bar{A}}, \tag{35a}$$

$$\beta_j = \beta_1 \Big/ \prod_{l=1}^{j-1} \Omega_{1l}^{A\bar{B}} . \tag{35b}$$

The general formulae presented above can be applied to the particular models presented earlier and to any other multiplicative models for the cross-classification table (e.g., the various models introduced in Goodman 1972).

Some of the results presented earlier in this note provide further insight into the meaning of some of the specific models considered herein. The general formulae presented above can be applied to a wide range of potentially useful specific models, and they will provide greater flexibility in the possible interpretations of these models. Let us turn now to a brief discussion of statistical techniques for the analysis of these models.

To estimate the expected frequencies under the "uniform association model" (as described in formulae [10]-[15], [26], [30a], or [30b]), Duncan (1979) used a trial-and-error search method in conjunction with the iterative procedure of the computer program ECTA (Everyman's Contingency Table

[30] The remark in n. 29 can be applied also to the products

$$\prod_{s=1}^{i-1} \Omega_{js}^{B\bar{A}} \quad \text{and} \quad \prod_{l=1}^{j-1} \Omega_{il}^{A\bar{B}}$$

in formulae (34b) and (34a). Formula (34a) calculates the expected count F_{ij} in terms of the parameters α_i (i.e., the row factors) and the set of odds $\Omega_{il}^{A\bar{B}}$, whereas formula (34b) calculates the F_{ij} in terms of the parameters β_j (i.e., the column factors) and the set of odds $\Omega_{js}^{B\bar{A}}$.

Analyzer).[31] This search method required the repeated calculation of "starting tables" for each successive application of the ECTA program. Instead of this trial-and-error search method, we can use a single iterative procedure that leads directly to the correct estimates.

For any model that can be described in terms of a tabular display that contains a multiplicative constant (say, δ or γ) but does not contain any higher powers of that constant (see, e.g., tables 5 and 6), the general iterative procedure presented in Goodman (1972) is available.[32] This iterative procedure can also be applied when the tabular display for the model contains various multiplicative constants (say, δ_1, δ_2, δ_3) but does not contain any higher powers of those constants (see, e.g., table 7). When the tabular display for the model contains a multiplicative constant (say, δ or ζ) and higher powers of that constant (see, e.g., tables 8, 10–13), the iterative scaling process in Darroch and Ratcliff (1972) is available.[33] This procedure can also be applied when the tabular display for the model contains various multiplicative constants (say, δ_1, δ_2, δ_3, δ_4) and higher powers of those constants (see, e.g., table 9).

The procedures referred to above can be applied to tabular displays for the full cross-classification table, and they can be modified in a straightforward way in order to make them applicable to tables in which some of the cells are omitted (see, e.g., Goodman 1968, 1972, 1978). For the sake of brevity, the present note will not go into these matters.

REFERENCES

Darroch, J. N., and D. Ratcliff 1972. "Generalized Iterative Scaling for Log-linear Models." *Annals of Mathematical Statistics* 43 (October): 1470–80.

Duncan, O. D. 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 84 (January): 793–803.

Goodman, L. A. 1965. "On the Statistical Analysis of Mobility Tables." *American Journal of Sociology* 70 (March): 564–85.

[31] The ECTA computer program, prepared by R Fay and the present author, is available from the author.

[32] The iterative scaling process in Darroch and Ratcliff (1972), and the modified Newton-Raphson method in Haberman (1974), could also be applied, but the procedure in Goodman (1972) is simpler when it is applicable. The cyclic iterative process in Darroch and Ratcliff (1972) reduces to the simpler iterative procedure in Goodman (1972) when the latter procedure is applicable. Although the modified Newton-Raphson method will usually require fewer iterations, the iterative procedure in Goodman (1972) is easier to apply; and even when the iterative procedure requires many iterations, the actual cost of using the procedure (i.e., the cost of a run applying a given computer program on an electronic computer) will usually be small. The modified Newton-Raphson method and the iterative scaling process in Darroch and Ratcliff (1972) can be used for all of the models presented in this note.

[33] In particular, the iterative scaling process is available for models (10)–(15). For these particular models and some other related models, an alternative iterative procedure was introduced and applied in Goodman (1978). This alternative procedure is relatively easy to apply, and it converges more rapidly than the Darroch-Ratcliff procedure in many cases.

———. 1968. "The Analysis of Cross-classified Data: Independence, Quasi-Independence, and Interactions in Contingency Tables with or without Missing Entries." *Journal of the American Statistical Association* 63 (December): 1091–131.

———. 1969. "How to Ransack Social Mobility Tables and Other Kinds of Cross-Classification Tables." *American Journal of Sociology* 75 (July): 1–39.

———. 1972. "Some Multiplicative Models for the Analysis of Cross-classified Data." Pp. 649–96 in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, edited by L. Le Cam, J. Neyman, and E. L. Scott. Berkeley: University of California Press.

———. 1978. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Row Categories and Ordered Column Categories." Unpublished paper.

Haberman, S. J. 1974. *The Analysis of Frequency Data*. Chicago: University of Chicago Press.

Hauser, R. M. 1978. "A Structural Model of the Mobility Table." *Social Forces* 56 (March): 919–53.

Simon, G. 1974. "Alternative Analyses for the Singly-ordered Contingency Table." *Journal of the American Statistical Association* 69 (December): 971–76

A CAUTIONARY ADDENDUM

Formulae (32) and (33) should remind the reader that we have considered here only models for F_{ij} that include α_i and β_j, only models for $\Omega_{ij}{}^{A\bar{B}}$ that include $(\beta_j/\beta_{j+1}) = \psi_j{}^B$, and only models for $\Omega_{ji}{}^{B\bar{A}}$ that include $(\alpha_i/\alpha_{i+1}) = \psi_i{}^A$.

Dynamic Analysis of Event Histories[1]

Nancy Brandon Tuma and Michael T. Hannan
Stanford University

Lyle P. Groeneveld
SRI International

There is wide interest among sociologists in the study of change but little reflection of this interest in sociological research methods. In this paper we consider the advantages of and procedures for dynamic analysis of event-history data—data giving the number, timing, and sequence of changes in a categorical dependent variable. We argue for grounding this analysis in a continuous-time stochastic model. This permits the data to be fully utilized; it also allows a unified treatment of the various outcomes analyzed in the many approaches that use only part of the information contained in such data. We focus on the familiar continuous-time Markov model, summarize its properties, report its implications for various outcomes, describe extensions to deal with population heterogeneity and time dependence, and outline a maximum-likelihood procedure for estimating the extended model from event-history data. The discussion is illustrated with an empirical analysis of the effects of an income-maintenance experiment on change in marital status. We conclude by contrasting event-history analysis with cross-sectional analysis, event-count analysis, and panel analysis. We find that event-history analysis has substantial advantages over the other approaches.

Though most sociologists profess an interest in social processes—how social behavior and social systems change over time—this interest is seldom reflected in their research. They usually examine relationships among phenomena at only one point in time. Even when temporal data are used, as in

[1] We gratefully acknowledge the research assistance of Helen Garrison, Beverly Lauwagie, Marlos Viana, and Barbara Warsavage, and the helpful comments of anonymous referees We appreciate the continuing support and encouragement of Robert G. Spiegelman, director of the Center for the Study of Welfare Policy, SRI International. Part of the analytic work reported herein was supported by the National Institute of Education Grant no NIE-G-76-0082 The rest of the analytic work and all empirical research reported herein was performed pursuant to contracts with the states of Washington and Colorado, prime contractors for the Department of Health, Education, and Welfare under contract nos. SRS-70-53 and HEW-100-89-0004, respectively. During the period of this research, Hannan was a fellow at the Center for Advanced Study in the Behavioral Sciences and was supported by National Science Foundation grant BNS76-22943-A01. The opinions expressed in the paper are those of the authors and should not be construed as representing the opinions or policies of the states of Washington or Colorado or any agency of the United States government.

experiments and panel studies, sociologists seldom study dynamics—the time paths of change. Instead, the focus is on change from one equilibrium level to another, as measured, for example, by pre- and postexperiment differences or by levels at successive waves of a panel.

To some extent professed interests and actual research diverge because data suitable for dynamic analysis are scarce. But opportunities to collect such data are often bypassed because investigators are uncertain how to utilize detailed information on change over time. We are struck particularly by the failure of social scientists to gather and analyze dynamically data on changes in categorical variables. We hope to stimulate interest in the collection and analysis of what we call event histories. An event history records dates of events that occur for some unit of analysis. Examples of events include changes in categorical variables describing individuals—such as marital status, employment status, health status, and membership in voluntary associations—as well as those applying to social collectivities—such as changes in political regimes and outbreaks of strikes, riots, and wars.

This paper has two main goals. We wish to show the value of dynamic analysis of event-history data for the sociological study of change in categorical variables. We also wish to describe in some detail a practical procedure for making full use of event-history data to answer the kinds of questions that interest social scientists.

Section I begins by noting the wide range of empirical analyses permitted by event-history data—including, of course, application of existing techniques for analyzing cross-sectional and panel data. We argue that the use of continuous-time stochastic models allows a unified treatment of the outcomes analyzed in these other approaches, and the full use of the information in event-history data. Thus we take the position that event-history analysis should be linked to an explicit mathematical model of the social process being studied. We focus on a familiar model, the continuous-time Markov model introduced into sociology by Coleman (1964). We summarize properties of the model and its implications for various outcomes. We also describe extensions of it to deal with population heterogeneity and time dependence; these extensions make the model much more useful for sociological research. Section I concludes with an outline of a maximum-likelihood procedure for using event-history data to estimate and test the extended model.

Section II gives an empirical illustration. This involves an elaboration of the analysis presented in Hannan, Tuma, and Groeneveld (1977) of the effects of an income-maintenance experiment on change in marital status. To illustrate the extension dealing with population heterogeneity, we report estimates of the effects of experimental treatments on rates of marriage, marital dissolution, and experimental attrition. To illustrate the extension dealing with time dependence, we report estimates of how treatment effects vary over the experimental period. We conclude by showing that the esti-

mated model not only provides much more information about the process than do other procedures more familiar to sociologists but also fits the data well—at least as well as the alternative procedures.

Finally, in Section III we contrast dynamic analysis of event-history data with other kinds of analysis. In particular, we compare it with cross-sectional analysis (assuming the process is in equilibrium), event-count analysis, and panel analysis. We find that when event-history data can be obtained, the strategy and procedures that we advocate have substantial advantages over these three alternatives.

A paper of this sort builds on a number of intellectual traditions and on contributions of previous authors. Our debt to classical mathematical sociology is obvious in our discussion of this model and its implications. Readers familiar with the causal-modeling (or structural equations) tradition should discern its influence on our approach to modeling causal structure and time dependence. In a companion paper (Tuma and Hannan 1978), we have sketched the genealogy of our statistical approach and have briefly reviewed some alternative estimation strategies.

I. CHOICE OF A MODEL

Any data set permits a variety of empirical analyses. When we analyze event histories, we have even more options than usual. This is seen clearly in the various dependent variables that can be selected from event histories. For example, consider how different investigators interested in marital stability might utilize marital histories to study changes in marital status. Some would study the probability of at least one marital dissolution within some fixed period. Others might study the average length of marriage. Others might choose to study the number of marital events during a fixed period. And some might analyze changes in the proportion of a population that is married.

What are we to make of the richness of the information contained in event histories? We certainly have no objection to different analysts focusing on different aspects of a problem. However, the various outcomes listed above are not independent. Moreover, none of them conveys all the information contained in an event history. We believe that much is gained from designing analyses that utilize all the information in an event history, that is, which use information on the number, sequence, and timing of events.

Our solution to this problem—not necessarily the only solution[2]—rests upon the recognition that we can infer each of the dependent variables listed earlier if we can describe each person's status at every moment in time.

[2] A simultaneous linear equations model might appear to be one alternative. However, the same exogenous variables appear in every equation, leading to underidentification of the structural parameters.

More generally, we can unify the analysis of various outcomes that can be extracted from event histories by building a model that describes the state of the categorical variable at every moment in time. This strategy has the advantage of parsimony—one model replaces several. It also has the advantage of dealing with the interdependence among the various outcomes. Of course, we expect to pay a price for such analytic power. In particular, we must make some restrictive assumptions about the process generating the events. Because of these assumptions, our model may fit sample data less well for any particular outcome at any arbitrary time than a model designed specifically to account for that outcome. After we have described our model and estimation procedure, we return to this issue as part of an extended empirical illustration of our methodology.

A Markov Model of Events

To use event histories fully, we must formulate an explicit model of the process generating the events. Because the model must account for states of categorical variables at every moment in time, and because the events can occur at any time, we choose continuous-time models. Exactly what type of model should be used depends on the substantive application; a variety of continuous-time models could be employed with the general methodology we propose. Here we concentrate on one type of model: a finite-state, continuous-time Markov model. We begin by stating briefly the formal assumptions of this model.[3]

Let $Y(t)$ be a random variable denoting the state of the categorical variable occupied by a member of a homogeneous population at time t. The set of all possible values of $Y(t)$ is called the state space of the process; this is assumed to be finite. Let $p_{jk}(u, t)$ represent the probability that someone in state j at time u is in state k at time t, $u < t$:

$$p_{jk}(u, t) = \Pr[Y(t) = k \mid Y(u) = j], \qquad (1)$$

and let $P(u, t)$ denote the matrix of these transition probabilities, $P(u, t) = \|p_{jk}(u, t)\|$. For example, suppose that $j = k = 2 =$ married; then $p_{22}(u, t)$ stands for the probability that those married at time u are also married at time t.

Next we define $r_{jk}(t)$ as the instantaneous rate of transition from state j to state k at time t. The transition rate $r_{jk}(t)$ is the limit, as Δt approaches zero, of the probability of a change from j to k between t and $t + \Delta t$, per unit of time:

$$r_{jk}(t) = \lim_{\Delta t \to 0} \frac{p_{jk}(t, t + \Delta t)}{\Delta t}, \quad j \neq k. \qquad (2)$$

[3] For a more thorough discussion, see Cox and Miller 1965, Feller 1968, Karlin and Taylor, 1975, or any other standard text on stochastic processes.

The rate of leaving state j at time t, $r_j(t)$, is as follows:

$$r_j(t) = \sum_{\substack{k \\ k \neq j}} r_{jk}(t)\,. \tag{3}$$

By the Markov assumption, the Chapman-Kolmogorov identity holds:

$$P(v, t) = P(v, u)P(u, t) \quad v \leq u \leq t\,. \tag{4}$$

Given this, and the usual additional assumptions (continuity, probabilities ranging between 0 and 1, etc.),[4] it can then be shown that

$$\frac{dP(u, t)}{dt} = P(u, t)R(t)\,, \tag{5}$$

where R is a matrix in which the j–kth off-diagonal element is the transition rate, $r_{jk}(t)$, and the jth diagonal element is the negative of the rate of leaving, $-r_j(t)$. In the time-independent case, $R(t) = R$, equation (5) has the solution:

$$P(u, t) = \sum_{m=0}^{\infty} \frac{R^m(t-u)^m}{m!} = e^{(t-u)R}\,, \tag{6}$$

where a matrix raised to the zero power is defined to equal the identity matrix I (so $R^0 = I$).

The first task in using such a model is specification of the state space, that is, the exhaustive and mutually exclusive set of discrete values of $Y(t)$. For example, in our analysis of the effects of income-maintenance treatments on marital stability (see Section II) the relevant states of $Y(t)$ are 1 (not married), 2 (married), and 3 (attrited). State 3 includes those who refuse to participate in the study, cannot be located, die, or emigrate outside the continental United States. Attrition is an absorbing state; it cannot be left. Thus R has the form:

$$R = \begin{Vmatrix} -r_1 & r_{12} & r_{13} \\ r_{21} & -r_2 & r_{23} \\ 0 & 0 & 0 \end{Vmatrix}. \tag{7}$$

Although the elements of $P(u, t)$ cannot be written as explicit functions of the transition rates for a general matrix R, they can be so expressed when R has the form shown in (7). In particular, it can be shown that $P(u, t)$ has the elements:[5]

$$p_{jj}(u, t) = \frac{1}{(\lambda_1 - \lambda_2)}\left[(r_k + \lambda_1)e^{\lambda_1(t-u)} - (r_k + \lambda_2)e^{\lambda_2(t-u)}\right], \tag{8}$$

[4] Mathematically these assumptions can be expressed as: $0 \leq p_{jk}(u, t) \leq 1$; $\Sigma_k\, p_{jk}(u, t) = 1$; $P(t, t) = I$

[5] Eqq. (8) through (12) do not apply to a general three-state model, i.e., one in which r_{31} and r_{32} are greater than zero.

$$p_{jk}(u, t) = \frac{r_{jk}}{(\lambda_1 - \lambda_2)} [e^{\lambda_1(t-u)} - e^{\lambda_2(t-u)}] , \tag{9}$$

$$p_{j3}(u, t) = 1 + \frac{1}{(\lambda_1 - \lambda_2)} [(r_{j3} + \lambda_2)e^{\lambda_1(t-u)} - (r_{j3} + \lambda_1)e^{\lambda_2(t-u)}] , \tag{10}$$

$$p_{31}(u, t) = p_{32}(u, t) = 0 , \tag{11}$$

$$p_{33}(u, t) = 1 , \tag{12}$$

where $j = 1$ or 2, $k = 3 - j$, $\lambda_1 \neq \lambda_2$, and

$$\lambda_1 = -[r_1 + r_2 + \sqrt{(r_1 - r_2)^2 + 4r_{12}r_{21}}\,]/2 , \tag{13}$$

$$\lambda_2 = -[r_1 + r_2 - \sqrt{(r_1 - r_2)^2 + 4r_{12}r_{21}}\,]/2 . \tag{14}$$

The explicit inclusion of attrition as a state is an important feature of our application of a Markov model. All too frequently investigators with temporal data analyze only the fraction of the original sample that has not been lost through attrition by the time of some later measurement. This procedure implicitly assumes that attrition has no impact on the assessment of the effects of causal variables. By explicitly including attrition as a state, we avoid this highly dubious assumption. However, it is still necessary to assume that the same model applies both to those lost through attrition and to those who are not. This assumption could be incorrect but cannot be tested without a follow-up study of marital status changes among those who did not remain in the active sample.

The inclusion of attrition as a state complicates considerably both the derivation of equations and our discussion. For this reason, parts of our discussion are based on the two-state model in which there is no attrition:

$$R = \left\| \begin{matrix} -r_1 & r_{12} \\ r_{21} & -r_2 \end{matrix} \right\| . \tag{15}$$

Then we obtain

$$p_{jk}(u, t) = 1 - p_{jj}(u, t) = \frac{r_{jk}}{r_{12} + r_{21}} [1 - e^{-(r_{12}+r_{21})(t-u)} \tag{16}$$

for $j, k = 1, 2$ and $j \neq k$.

Implications of the Model

We noted earlier that use of such a continuous-time stochastic model lets us derive implications about a variety of observable variables. In this section we show this by discussing a number of well-known results about Markov models.

Time between transitions.—It is widely known that for a Markov model

the length of time between transitions has an exponential distribution whose parameter depends on the transition rates (see, e.g., Breiman 1969, chap. 7). In particular, let $F_j(t|u)$ represent the probability of a transition from state j before time t, given that state j is occupied at time u. $F_j(t|u)$ is usually called the cumulative probability distribution function or simply the probability distribution function. We can show that for models in which $R(t) = R$,

$$F_j(t|u) = 1 - e^{-(t-u)r_j} . \tag{17}$$

Later we use the definition that

$$G_j(t|u) = 1 - F_j(t|u) . \tag{18}$$

$G_j(t|u)$ is often called the survivor function because it gives the probability that a unit in state j at time u remains (or survives) in state j until time t. For models in which $R(t) = R$,

$$G_j(t|u) = e^{-(t-u)r_j} = e^{-(t-u)[\Sigma_k r_{jk}]} . \tag{19}$$

This says, for example, that the probability that a marriage existing at time u survives until a later time t declines exponentially as the length of the interval $(t - u)$ increases. This monotonic decline occurs because it becomes increasingly likely as time t increases that either the marriage breaks up or the couple drops from the experiment.

We also use the probability density function:

$$f_j(t|u) = \frac{dF_j(t|u)}{dt} = \frac{-dG_j(t|u)}{dt} . \tag{20}$$

It can be shown that in general

$$f_j(t|u) = r_j(t)G_j(t|u) . \tag{21}$$

For the particular case in which $R(t) = R$, that is, in which the transition rates are constant over time,

$$f_j(t|u) = r_j e^{-(t-u)r_j} . \tag{22}$$

The probability of a change from state j to some other state between t and $(t + dt)$ is approximately equal to $f_j(t|u)dt$. Equation (22) shows that this probability initially (i.e., at $[t - u] = 0$) equals r_j and declines exponentially as the length of the interval $(t - u)$ increases. In other words, the probability of leaving a state varies over time even when transition rates are constant. This is one of the main advantages of modeling social processes in terms of transition rates instead of probabilities of change.

The average duration of state occupancies (e.g., expected duration of marriages and of the intervals between marriages) when $R(t) = R$ is easily shown to be

$$E(t - u) = 1/r_j , \tag{23}$$

where u and t denote the times of entry and exit, respectively, from state j.

Given that a change occurs at time t, r_{jk}/r_j is the conditional probability that k is the destination. Thus, with this model, we can account for both the probability of at least one change within any specified time interval and the conditional probability that a change involves movement into some specified state.

State probabilities.—We can also use the model to find the unconditional probability of being in any specified state at any point in time (e.g., the probability of being married at any specified time). Let $p_k(t)$ denote the unconditional probability of being in state k at time t, and $p(t)$ represent a row vector giving the probability distribution among the states at time t. It is easy to see that

$$p_k(t) = \sum_j p_j(u) p_{jk}(u, t) , \tag{24}$$

or

$$p(t) = p(u) P(u, t) . \tag{25}$$

Usually we are interested in $p(t)$ when $p(u) = p(0)$, the distribution at the start of the process.

For simplicity we begin by considering the two-state model, for example, for the study of marital stability when there is no attrition. Inserting equation (16) into (24), we obtain:

$$p_1(t) = p_1(0)e^{-(r_{12}+r_{21})t} + \frac{r_{21}}{r_{12} + r_{21}} [1 - e^{-(r_{12}+r_{21})t}] , \tag{26}$$

$$p_2(t) = p_2(0)e^{-(r_{12}+r_{21})t} + \frac{r_{12}}{r_{12} + r_{21}} [1 - e^{-(r_{12}+r_{21})t}] . \tag{27}$$

As time t becomes very large, $p_j(t)$ approaches a steady-state (or equilibrium) value: $p_1(\infty) = r_{21}/(r_{12} + r_{21})$, $p_2(\infty) = r_{12}/(r_{12} + r_{21})$. So $p_j(t)$ is a weighted average of the initial proportion in state j and the steady-state probability in state j. The weight given to $p_j(0)$ declines exponentially over time, while the weight given to the steady-state proportion in state j increases over time until it reaches one.

For the three-state model, for example, for the study of marital stability when there is attrition, we find (by substituting equations [8] through [12] into [24]) that:

$$\begin{aligned} p_1(t) = {} & (e^{\lambda_2 t}\{(r_2 + \lambda_2)p_1(0) + r_{21}[1 - p_1(0)]\} \\ & - e^{\lambda_1 t}\{(r_2 + \lambda_1)p_1(0) + r_{21}[1 - p_1(0)]\})/(\lambda_2 - \lambda_1) , \end{aligned} \tag{28}$$

$$\begin{aligned} p_2(t) = {} & (e^{\lambda_2 t}\{(r_1 + \lambda_2)[1 - p_1(0)] + r_{12}p_1(0)\} \\ & - e^{\lambda_1 t}\{(r_1 + \lambda_1)[1 - p_1(0)] + r_{12}p_1(0)\})/(\lambda_2 - \lambda_1) , \end{aligned} \tag{29}$$

$$p_3(t) = 1 + (e^{\lambda_2 t}\{\lambda_1 + r_{13}p_1(0) + r_{23}[1 - p_1(0)]\} - e^{\lambda_1 t}\{\lambda_2 + r_{13}p_1(0) + r_{23}[1 - p_1(0)]\})/(\lambda_2 - \lambda_1)\,, \tag{30}$$

where λ_1 and λ_2 are given by (13) and (14), respectively. It is important to emphasize that $r_1 = r_{12} + r_{13}$, $r_2 = r_{21} + r_{23}$ and that λ_1 and λ_2 depend on r_1 and r_2. Thus each of the above equations depends on the attrition rates, r_{13} and r_{23}.

Expected number of events.—Finally, we can derive relationships between the fundamental parameters of the model and the expected number of events that occur in any interval of time. This requires extension of the ordinary Poisson model of the number of events in a time interval, since the Poisson model assumes that only one kind of event can occur. Let $n_{jk}(0, t)$ denote the number of changes to state k between times 0 and t when state j is occupied at time 0, $0 \leq t$. For example, $n_{22}(0, t)$ stands for the number of marriages formed before t for a person who is married at time 0. It can be shown that for the two-state model (in which there is no attrition):

$$E[n_{jj}(0, t)] = \frac{r_{12}r_{21}t}{(r_{12} + r_{21})} + \frac{r_{12}r_{21}}{(r_{12} + r_{21})^2}[e^{-t(r_{12}+r_{21})} - 1]\,, \tag{31}$$

$$E[n_{jk}(0, t)] = \frac{r_{12}r_{21}t}{(r_{12} + r_{21})} - \frac{r_{jk}^2}{(r_{12} + r_{21})^2}[e^{-t(r_{12}+r_{21})} - 1]\,, \tag{32}$$

where $j, k = 1, 2$ and $j \neq k$. For the three-state model we find that:

$$E[n_{jj}(0, t)] = \frac{r_{12}r_{21}}{(\lambda_1 - \lambda_2)}\left[\frac{(1 - e^{\lambda_1 t})}{\lambda_1} - \frac{(1 - e^{\lambda_2 t})}{\lambda_2}\right], \tag{33}$$

$$E[n_{jk}(0, t)] = \frac{r_{jk}}{(\lambda_1 - \lambda_2)}\left[\frac{(r_k + \lambda_2)(1 - e^{\lambda_2 t})}{\lambda_2} - \frac{(r_k + \lambda_1)(1 - e^{\lambda_1 t})}{\lambda_1}\right], \tag{34}$$

$$E[n_{j3}(0, t)] = \frac{1}{(\lambda_1 - \lambda_2)}\left[\frac{(r_{jk}r_{k3} + r_{j3}r_k + r_{j3}\lambda_2)(1 - e^{\lambda_2 t})}{\lambda_2} - \frac{(r_{jk}r_{k3} + r_{j3}r_k + r_{j3}\lambda_1)(1 - e^{\lambda_1 t})}{\lambda_1}\right], \tag{35}$$

where $j, k = 1, 2$ and $j \neq k$.

The point we wish to emphasize is that these diverse observable measures of event histories are all explicit functions of the rates that define the Markov model. It is not necessary that each of these relationships be estimated separately. Any one of them provides an approach to estimation. Once the rates have been estimated, the remaining quantities can be calculated using these functions. The empirical analysis in Section II uses them in evaluating the fit of our model.

Two Extensions of the Model

Markov models have been used quite often by sociologists in the past 20 years or so. In the simple form discussed above, these models have invariably failed to fit the data. Sociologists' efforts to increase realism and improve fit have concentrated on modification of two assumptions of the simple Markov model: time-dependence (see, e.g., Mayer 1972; Sørensen 1975; Tuma 1976) and population homogeneity (see, e.g., Blumen, Kogan, and McCarthy 1955; Coleman 1964, 1973; Land 1971; McFarland 1970; Spilerman 1972*a*, 1972*b*; Tuma 1976). These sorts of extensions are straightforward in the strategy we propose.

Population heterogeneity.—To this point we have assumed that the units studied are homogeneous, in other words, that their behavior is governed by the same, constant transition rates. Instead, following Coleman (1964, 1973) and Tuma (1976), we may assume that the same, constant rates only govern the behavior of units with identical values on a set of observable, exogenous variables. In other words, we may establish relationships between observable variables, denoted by x, and the rates of change in categorical variables. The main sociological interest in most analyses of event histories lies in testing hypotheses concerning the effects of exogenous variables on rates. For example, in analyses of marital histories produced by the income-maintenance experiments, we concentrate on testing hypotheses about effects of experimental treatments on rates of marital dissolution and remarriage.

To introduce such causal relationships, we must state an explicit dependence of the unobservable rates on the observable variables. In our empirical analysis, we assume a log-linear relationship between each transition rate and x:[6]

$$\ln r_{jk} = \theta_{jk} x \,, \qquad \text{for all } j \text{ and } k, \; j \neq k \,, \tag{36}$$

where θ_{jk} represents a vector of parameters to be estimated. The log-linear relationship constrains r_{jk} to be positive for each individual, whatever the value of x, in accord with equation (2); we find that it usually fits data better than a linear relationship. We also assume that r_{jk} is finite for all j and k and for all individuals.

Time-dependence.—We have assumed to this point that the rates are con-

[6] As mentioned by an anonymous referee, we could have chosen to model r_j and m_{jk} ($= r_{jk}/r_j$) rather than r_{jk} as functions of exogenous variables. The choice involves a decision about the substantive nature of the process; it is not just a methodological issue. The approach suggested by the referee, which has also been advocated by Tuma (1976) and by Singer and Spilerman (1974), is appropriate when the decision to leave the current state and the choice of a destination are separate or sequential In the example used in this paper, it seems reasonable to think that the decision to leave a state and the choice of a destination are not separate; therefore, we have not adopted this other approach. For other social processes it may, indeed, be preferable to model r_j and m_{jk} separately. The method of estimation and accompanying software described below can readily be used to estimate $r_j = e^{ax}$ from data on the dates of entering and leaving state j

stant over time. Sociologists are accustomed to thinking about population heterogeneity (i.e., causal relationships) but have not devoted much attention to issues involving time dependence of effects, that is, the manner in which causal relationships change over time. However, sometimes theory indicates that rates of change are some specific function of time. For example, Sørensen (1975) has argued that rates of leaving a job are exponentially declining functions of labor-force experience, while Tuma (1976) has specified a model in which rates of job leaving are second-order polynomials in duration in the job. Such parametric forms of time dependence may be accommodated without difficulty in the strategy we discuss.

Sometimes we do not have any a priori hypothesis concerning time dependence. In an experiment, we might expect transition rates during an initial adjustment period to differ from rates later on but be unsure whether initial rates should be larger or smaller than later rates. In these circumstances we can define a set of time periods and allow the effects of experimental variables to vary freely from period to period while remaining constant within each period. In this situation the time periods may be arbitrary, a disadvantage of this approach. However, in other situations, we may have some idea about points on the time axis when rates may change in some way. For example, rates of collective violence may change when there is a change in the political regime or during periods of warfare.

Mathematically stated, a model in which rates vary from one time period to another is:

$$\ln r_{jkp} = \boldsymbol{\theta}_{jkp} \mathbf{z}_p \,, \qquad \tau_{p-1} \leq t < \tau_p \,, \tag{37}$$

where τ_p is the last moment in period p, $\tau_0 = 0$ is the starting time, $\mathbf{z}_p$ is a vector of exogenous variables affecting the rate in period p, and $\boldsymbol{\theta}_{jkp}$ is a vector of parameters giving effects on r_{jkp} in period p. We illustrate the use of a model similar to this one below.

Estimation and Testing

The linear structural equations models with which sociologists have become quite familiar in recent years have made common the view that the estimation equation is identical to the model. But this is not necessarily the case. The equations relating observable outcomes (average duration, number of events, etc.) to the rates are all possible candidates for estimation equations, but they are not equally promising because they differ in complexity.

Building upon the work of Bartholomew (1957) and Albert (1962), we form maximum-likelihood estimators of rates (and of causal effects on rates) using data on the dates and kind of events. This approach offers a number of advantages in the present circumstances. First, maximum-likelihood estimators have good properties for large samples under fairly general conditions. Second, maximum-likelihood estimators retain their good properties under

any monotonic transformation. Thus one can use maximum-likelihood estimators of rates to form maximum-likelihood estimators for expected durations and other monotonic functions of rates. Third, maximum-likelihood procedures permit a satisfactory solution of what is called the censoring problem. Data collected during an observation period too short to produce a change for every case are said to be censored, and errors of inference are likely if appropriate measures for dealing with censoring are not adopted (Tuma and Hannan 1978).

Assuming independent observations on the N different cases (units) being analyzed, we can write a likelihood function that uses all information in event histories—namely, the number, timing, and sequence of events. This information can be represented by the following kinds of variables: w_{mi} (a dummy variable that equals unity if case i's mth event occurs in the observation period), t_{mi} (the time of case i's mth event if $w_{mi} = 1$ and the time that case i is last observed if $w_{mi} = 0$), and v_{mji} (a dummy variable that equals unity if case i's mth event is observed and consists of a change to state j and otherwise equals zero). We define t_{0i} as the start of the observation period on case i and v_{0ji} as unity if state j is occupied by i at t_{0i} and otherwise as zero. Note that

$$\sum_{m=1}^{\infty} w_{mi}$$

equals the total *number* of events observed to occur to case i, while the t_{mi}'s are the *times* at which these events occur and the v_{mji}'s indicate the *sequence* of events.

For simplicity we begin with the likelihood equation that uses information on the first event only:

$$\mathcal{L} = \prod_{i=1}^{N} \prod_{j=1}^{n} \Big\{ [G_j(t_1 | t_0, x)]^{(1-w_1)v_{0j}} \cdot [f_j(t_1 | t_0, x)]^{w_1 v_{0j}} \cdot \prod_{k=1}^{n} [r_{jk}(x)/r_j(x)]^{-w_1 v_{0j} v_{1k}} \Big\}, \tag{38}$$

where N is the number of cases, n is the number of states, and the subscript i on variables (namely, w_{mi}, t_{mi}, v_{mji}, and x_i) has been suppressed for clarity. The first term (in brackets), the survivor function, gives the probability that the first event has not occurred by time t_1 (see eq. [18]). Since any number raised to the zero power equals unity, the first term differs from unity only for those cases that have *not* experienced a first event ($w_1 = 0$) *and* are in state j at time t_0 ($v_{0j} = 1$). The second term is the probability density that the first event occurs at time t_1 (see eq. [20]). This term differs from unity only for those cases that have experienced at least one event ($w_1 = 1$) and

are in state j at time t_0 ($v_{0j} = 1$). The third term is the probability that the first event consists of a move to state k, given that state j is left. This term differs from unity only for those cases that have experienced at least one event ($w_1 = 1$), that are in state j at time t_0 ($v_{0j} = 1$), and whose first event consists of a move to state k ($v_{1k} = 1$).

Equation (38) can be written more simply by substituting (21) into (38) and by then collecting terms:

$$\mathcal{L} = \prod_{i=1}^{N} \prod_{j=1}^{n} \left\{ [G_j(t_1 | t_0, x)]^{v_{0j}} \cdot \prod_{k=1}^{n} [r_{jk}(x)]^{w_1 v_{0j} v_{1k}} \right\} . \qquad (39)$$

When $R(t) = R$, so that (19) holds, we can simplify this still more:

$$\mathcal{L} = \prod_{i=1}^{N} \prod_{j=1}^{n} \left\{ \prod_{k=1}^{n} [e^{-(t_1 - t_0) r_{jk}(x)}]^{v_{0j}} \right\} \cdot \left\{ \prod_{k=1}^{n} [r_{jk}(x)]^{w_1 v_{0j} v_{1k}} \right\} ,$$

$$\mathcal{L} = \prod_{j=1}^{n} \prod_{k=1}^{n} \left\{ \prod_{i=1}^{N} [e^{-(t_1 - t_0) r_{jk}(x)}]^{v_{0j}} \cdot [r_{jk}(x)]^{w_1 v_{0j} v_{1k}} \right\} . \qquad (40)$$

Note that the term in braces in (40) above does not depend on $r_{j'k'}$ where $j' \neq j$ and $k' \neq k$. Therefore, the maximum-likelihood estimates of a given r_{jk}—or of the effects of causal variables on r_{jk} as in (36)—can be obtained by maximizing the term in braces for that particular j and k. This fact is quite important in practice. It means that we can estimate selected transition rates without having data on *all* kinds of transitions. To estimate r_{jk}, we need data only on the times of all first events (t_1) for cases originally in state j ($v_{0j} = 1$) and on the outcome of first events (the value of v_{1k}). This means we can concentrate data collection and analysis on units originally in states of particular theoretical interest and can ignore events occurring to other units—unless we wish to model or predict the overall evolution of the entire process.

Equations (38) through (40) can be generalized in a straightforward way to the situation in which data on all observed events (and not just the first) are used in the estimation procedure. We give the equations corresponding to (38) and (40) only, leaving the intermediary equation to be supplied by the interested reader:

$$\mathcal{L} = \prod_{i=1}^{N} \prod_{m=1}^{\infty} \prod_{j=1}^{n} \left\{ [G_j(t_m | t_{m-1}, x)]^{(1-w_m) v_{m-1,j}} \cdot [f_j(t_m | t_{m-1}, x)]^{w_m v_{m-1,j}} \cdot \prod_{k=1}^{n} [r_{jk}(x)/r_j(x)]^{w_m v_{m-1,j} v_{mk}} \right\} . \qquad (41)$$

When $R(t) = R$, this simplifies to:

$$\mathcal{L} = \prod_{j=1}^{n} \prod_{k=1}^{n} \prod_{m=1}^{\infty} \prod_{i=1}^{N} \left\{ [e^{-(t_m - t_{m-1}) r_{jk}(X)}]^{v_{m-1,j}} \cdot [r_{jk}(x)]^{w_m v_{m-1,j} v_{mk}} \right\} . \quad (42)$$

Maximum-likelihood estimators of the parameters are found by maximizing $\mathcal{L}$ (or its logarithm).[7] The optimal asymptotic properties of maximum-likelihood estimators (consistency, asymptotic normality) are well known (e.g., see Dhrymes 1970). There is also evidence (Tuma and Hannan 1978) that the properties of the estimators obtained from (40) remain quite good even in small samples and with a high degree of censoring (i.e., when the mean of number of events is low).

A likelihood ratio test can be used to test nested models. Let $\mathcal{L}_\Omega$ be the likelihood for a model Ω with $q + s$ estimated parameters and $\mathcal{L}_\omega$ be the likelihood for the nested model ω that has s parameters constrained (usually to equal zero) and q parameters estimated. The likelihood ratio λ is defined to equal [max $(\mathcal{L}_\omega)$/max $(\mathcal{L}_\Omega)$]. Asymptotically $-2 \ln \lambda$ has a χ^2 distribution with q degrees of freedom, permitting us to test the fit of the model ω relative to the model Ω. Furthermore, it is possible to perform tests on the coefficients of individual variables using the estimated covariance matrix of the parameters (the inverse of the matrix of second derivatives of the natural logarithm of $\mathcal{L}$ with respect to the parameters). The square root of the variance of a parameter gives its estimated standard error, which can then be used in a standard fashion to calculate a t-statistic (or F-ratio) that tests whether the parameter differs from its value in the null hypothesis.

II. AN APPLICATION

To illustrate the models and methods discussed above we apply them to data on white women during the first two years of the Seattle and Denver Income-Maintenance Experiments (SIME/DIME). We combine data from Seattle and Denver.[8]

We have discussed the experimental design, nature of the sample, definitions of variables, and so forth earlier (Hannan, Tuma, and Groeneveld 1976, 1977). Here we use similar data that cover a period of 24 months instead of 18. We focus on estimating the effects of the three levels of income support (or guarantee) and the length of the treatment (three or five years). Because the design involved stratified random assignment to treatments, we include a number of preexperimental variables in the models: dummy variables representing six normal income categories, a site dummy

[7] A FORTRAN computer program called RATE has been developed to find the maximum-likelihood estimates of parameters in eqq. (36), (37), and (43), among others. Written documentation, test data, and test output are also available. For information, write the first author.

[8] For our definitions of marriage and marital dissolution, see Hannan et al. 1977.

(1 = Denver), a dummy variable for previous AFDC experience, number of children, a dummy variable for having any children under six, the woman's age, her wage rate, and her years of schooling.

Though other causal variables, as well as the experimental treatments, may have different effects on rates of events in different time periods, we are interested primarily in the time dependence of treatment effects. There are a variety of reasons for expecting the effects of experimental treatments to vary over the experimental period. Reasons for treatment effects to be smaller initially than later on include the possibilities that subjects do not fully understand the treatments at first and that they may need to search for an opportunity to change state (e.g., to find a marital partner). Reasons for treatment effects to be larger initially than later on include the possibility that the treatments become less salient as the end of the experiment approaches and the possibility that the treatments improve opportunities for a change for some fraction of subjects on the verge of changing marital status before the experiment. Indirect treatment effects (e.g., effects of treatments on work behavior, which in turn affect marital status) also can cause treatment effects to vary over time, but it cannot be predicted a priori whether indirect effects enhance or dampen the initial response. Our inability to predict the shape of the pattern of time variation is, of course, the reason for subdividing the observation period and letting rates vary from one period to another.

In the time-dependent models we began by using four time periods with end points of 0.5, 1.0, 1.5, and 2.0 years after the start of the experiment. These represented a compromise between two conflicting goals. First, we wanted the number of observed events per period to be large so that the standard errors of parameters would be comparatively small. Second, we wanted to have a large number of time periods so that we could detect the shape of the pattern of time variation (which according to our reasoning might decrease over time, rise over time, or rise and then decrease, etc.)

To improve the efficiency of estimation of effects still further, we estimated an equation that allowed treatment effects, but not the effects of other causal variables, to vary over time. This model can be represented as follows:

$$\ln r_{jkp} = \phi_{jk} x + \theta_{jkp} z_p , \qquad \tau_{p-1} \leq t < \tau_p , \tag{43}$$

where x is the vector of other causal variables, z_p is a vector of experimental treatment variables, and p refers to one of P periods.

There is a constant term in ϕ_{jk} and each θ_{jkp}; however, for each transition j to k, only P of these $P+1$ constants can be identified. Therefore, to achieve identification we arbitrarily constrain one of them to be zero.[9] The

[9] The same problem arises in regression analysis when dummy variables are used to represent categorical variables. One category must be omitted to achieve identification; the one chosen affects the interpretation of coefficients of the included dummy variables but does not affect predictions of the dependent variable.

P constants permit the rate of the control group to vary from one time period to another, even though the effects of other causal variables may not.[10]

The method of maximum likelihood can be used to estimate parameters in equation (43). The likelihood equation resembles (39) except that there is a different expression for the survivor function and each transition rate has an additional subscript p to indicate the time period to which it applies (for details, see the Appendix). In other respects procedures for estimating and testing the model and the effects of individual variables are the same as for a model with transition rates that are constant over time.

Results of the estimation could be reported in many different ways. Because the effect of a variable on the rate itself is of more interest than the effect on the log of the rate, we report the antilog of estimates $\hat{\theta}$. The antilog indicates the multiplier of the rate for a unit increase in a variable. For dummy variables, which we use to represent experimental treatments, the antilog of the coefficient of the variable is the ratio of the rate for those whose value on the dummy variable is unity to the rate for those in the omitted category. For example, if for some experimental treatment, $e^{\theta} = 2$, where θ is the coefficient of the dummy variable representing the experimental treatment, then the rate for those on the treatment is twice the rate for those in the control group. The percentage change in the rate for an experimental treatment relative to the rate for those in the control group is just $100\,(e^{\theta} - 1)$. Thus, if $e^{\theta} = 2$, the percentage change in the rate for this treatment relative to the rate for the control group is 100%.

Results for Time-independent Models

Table 1 gives the results for time-independent models of rates of marital status change and attrition. We have discussed similar results and their interpretation at length previously (Hannan et al. 1977, Tuma, Groeneveld, and Hannan 1976).[11] Here we comment briefly, concentrating on the effects of the support levels.

All four models improve significantly (at the .001 level) on a constant rate model. Inclusion of the set of experimental treatments improves significantly on a model that includes only the other causal variables in the case of the marital dissolution rate (.001 level), the attrition rate of married women (.10 level), and the attrition rate of unmarried women (.05 level), but not in the case of the remarriage rate.

Women in each support group have higher marital dissolution rates than women in the control group who are comparable in terms of values of other causal variables. Effects of the support levels are large (ranging from a

[10] The rate of the control group may vary over time because of aging, secular trends, etc.

[11] Although these papers do not explicitly mention attrition as a third state, the actual estimation procedures were the same as those reported here.

TABLE 1

EFFECTS OF VARIABLES ON RATES OF MARITAL STATUS CHANGE AND ATTRITION OF WHITE WOMEN: TIME-INDEPENDENT MODEL[a]

Variables	Marital Dissolution	Attrition of Married Women	Marital Formation	Attrition of Unmarried Women
Support level $3,800	2 29***	1 24	1.27	86
Support level $4,800	2.06***	96	1 08	.40**
Support level $5,600	1 57	53	.80	.64
Three-year program	86	79	1 01	91
Normal income levels ($):				
0–999	3 80***	2 52*	31**	82
1,000–2,999	2 71***	1 57	48	.64
3,000–4,999	2 21***	89	53	.51
5,000–6,999	1 88**	1 09	.54	.60
7,000–8,999	1 21	1 06	62	.65
Unclassified	4 13***	1 26	.64	1.12
One if on AFDC before enrollment	1 54**	85	1 09	.81
One if any children under 6 yr.	97	1.10	83	.62*
Children (N)	88**	.93	1.20**	1.01
Woman's age (yr.)	97***	1.00	.91***	95***
Woman's education (yr.)	.97	.95	1 02	98
Woman's wage ($/hr)	1 33**	1 41**	1.05	1.00
One if Denver	90	1 85***	92	1.79**
Constant	10***	.04***	2.60	.73
N	1376	1376	914	914
Likelihood ratio test for model (17 df)	107 55***	26.52*	97 87***	32.16**
Likelihood ratio test for experimental effects (4 df)	18 78***	8.03*	2 91	9 54**

[a] Coefficients are the multipliers of the rate for a unit change in a variable, i.e , exp (θ). A coefficient of 1 0 means the variable has no effect.

* $10 \geq P > .05$.

** $.05 \geq P > .01$.

*** $01 \geq P$.

57% to 129% increase in the rate relative to comparable controls) and statistically significant at the .01 level, except for the highest support.

As expected, attrition rates among those with a financial treatment are lower than those among comparable controls, except in the case of married women on the low support. However, there is no clear pattern to the effects of the different support levels on attrition, and none of the individual coefficients is significant.

Results for the Time-dependent Models

We have estimated time-dependent (four-period) models of rates of dissolution, remarriage, and attrition (see Tuma, Hannan, and Groeneveld [1977] for a more detailed discussion). The results of these analyses indicated that experimental effects on attrition rates did not vary significantly over time or in any patterned way. Experimental effects on the remarriage rate were significant at the .10 level, but there was no particular pattern to these effects, suggesting that they resulted from chance alone. However, experimental effects on the dissolution rate had a striking pattern, so we focus on these.

Table 2 shows results of the experimental effects on the dissolution rate when the total observation period (24 months) is treated as a single period, four six-month periods, and two periods (the first six months and the remaining 18 months). In all three models inclusion of the set of experimental treatments improves significantly on a model that contains only the other causal variables.

Support level effects over the four periods are plotted in figure 1. The plot shows that all three support levels produce an exceptionally large increase in the rate during the first half-year of the experiment. Except for the $5,600 support level in the second half-year, effects of the support levels in periods two through four are positive and show no clear pattern of variation over time. This suggests that (1) the support levels have transitory effects on dissolution rates of white women that subside after six months, but (2) they also have nontransitory positive effects on the dissolution rate. Furthermore, the effect of each support level in the first six-month period relative to its effect in each subsequent period is about the same for all three support levels. Although there is a clear pattern of time-varying effects, the four-period model does not improve significantly on the one-period model.

With these findings in mind, we estimated a two-period model in which the first half-year is distinguished from the rest of the experimental period. Relative rates across support levels were constrained to be equal in the two periods (see table 2), but the rate relative to controls was allowed to vary from one period to the other. Treatment effects in this model vary significantly (.05 level) over time, according to both the likelihood ratio test and

TABLE 2

TIME-DEPENDENCE OF EFFECTS OF TREATMENTS ON MARITAL DISSOLUTION RATES OF WHITE WOMEN[a] (N = 1,376)

Periods (Years)	Support $3,800	Support $4,800	Support $5,600	Three-Year Treatment	Financial Treatment	Likelihood Ratio Test (χ^2) for Effects of Experimental Treatments	Degrees of Freedom	Likelihood Ratio Test (χ^2) for Time-dependent Effects of Experimental Treatments	Degrees of Freedom
One-period model:									
(0–2.0) . .	2.29***	2 06***	1 57	86	. .	18 78***	4	. .	. . .
Four-period model:									
First period (0–0 5)	5.69***	3.37**	4.27**	.74	.				
Second period (0 5–1 0)	1 83	1 55	71	.76		33.71***	16	14.80	12
Third period (1.0–1.5)	2 07*	1 75*	1 34	.68	.				
Fourth period (1.5–2 0).	1 53	2 23*	1 50	1 47	.				
Two-period model:									
First period (0–0 5) .	1.96***	1 76**	1 35	.86	2 32**	23.33***	5	4.40**	1
Second period (0 5–2 0) . .	1.96***	1.76**	1 35	86	.				

[a] All equations contain the other causal variables given in table 1. Coefficients are exp ($\hat{\theta}$) and indicate the multipliers of the rate. A coefficient of 1.0 means the variable has no effect.

* .10 ≥ P > 0 5.

** .05 ≥ P > 0 01.

*** .01 ≥ P.

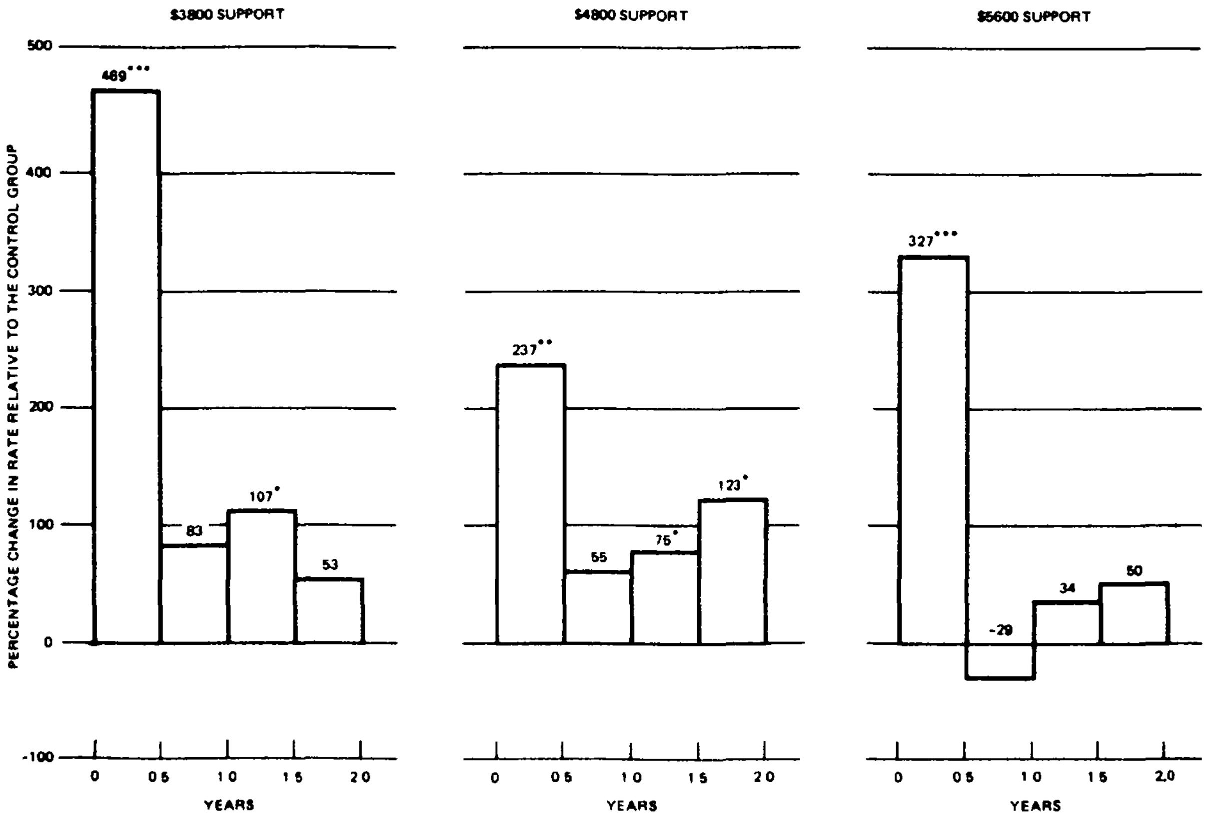


FIG. 1.—Relationship of income-maintenance support levels to marital dissolution rates of white women over time. *.10 ≥ P>.05, ** .05 ≥ P>.01, ***.01 ≥ P.

the F-test on a dummy variable for an effect of financial treatments in the first half-year.

The estimates for the two-period model (table 2) indicate that the effect of each support level is 2.32 times as large in the first half-year as it is thereafter. This is consistent with the possibility mentioned above that the treatments immediately changed the opportunities of some respondents on the verge of dissolving their marriages. But the treatments also seem to have changed the long-range opportunity structure. The effects of the support levels during the 0.5–2.0-year period are positive and significant (except for the $5,600 support), though 14%–15% lower than in the results for the time-independent model. Since the long-range impact of an income-maintenance program should not depend on transitory effects on rates during an initial adjustment period, the effects of the support levels in the 0.5–2.0-year period should provide more reliable estimates of the ultimate effects of income maintenance than estimates based on the one-period model.

How Well Does the Model Fit?

Markov models have a reputation for fitting data poorly. But our extensions of the Markov model should have helped to improve the ability of the model to fit the data. So far we have used likelihood ratio tests to assess the relative fit of a series of nested models. We have learned that some models do not improve upon others. Here we look at the absolute fit of the model and compare it with a common alternative.

Three main questions are involved in assessing our dynamic model: (1) To what extent do predictions based on our model differ systematically from observed values? (2) If there are systematic differences, are they related to the experimental treatments, causing us to make erroneous inferences about experimental effects on a particular outcome? (3) As compared with approaches that seek to minimize prediction errors for a single observable variable, how well does our model explain sample variation in outcomes related to marital stability?

In answering these questions we consider three kinds of outcomes: the probability of being in a given state (e.g., single) at any moment, the expected number of marriages and marital dissolutions in any given time interval, and the probability of leaving the original marital status in some time period. Our model implies that each of these is a function of marriage, dissolution, and attrition rates, as indicated in Section I. Thus for any woman we can use her support level and her values of other causal variables and the estimated effects of these variables to predict her rates of marriage, marital dissolution, and attrition. From the predicted rates we can then predict the kinds of outcome listed above. We chose to consider these outcomes at two arbitrary times: one and two years after the start of the experi-

ment. For each woman used in our analysis of transition rates, we retrieved the observed outcomes listed in table 3. These variables were not used directly in estimating transition rates; nor, of course, are they independent of those data, which is the reason for having a single model. We also predicted each woman's outcomes using the estimated effects from the one-period models of remarriage and attrition and the two-period model of marital dissolution (see tables 1 and 2).

To detect systematic differences between predictions and observed values we report the mean residual for each outcome, that is, the mean difference between observed and predicted variables. We also report the observed mean of each outcome because the relative size of a systematic difference is of some interest too. With predictions from a linear regression model, the mean residual would be zero. This need not be the case with predictions from our model. The results in table 3 show that the mean residuals for our predictions are usually small in both absolute (the largest is .02) and relative terms. There is little indication of any overall pattern in these differences, except for the last four dummy variables, which have consistently positive mean residuals. It is well known that no change in status has tended to be underpredicted in sociological applications of Markov models (e.g., see Blumen et al. 1955). Our introduction of population heterogeneity has made this a comparatively small problem, but it has not erased it entirely.

Small mean residuals could hide systematic differences associated with different treatments, which is clearly undesirable if differences between outcomes for controls and for those on financial treatments are of interest. To answer the second question, we performed one-way ANOVA on the residuals for each outcome. Treatment differences in the residuals never even approached statistical significance. (The smallest probability was greater than .50.)

In addressing the third question, we focus on a single, common, inexpensive alternative—linear regression analysis. We regressed each observed outcome on the prediction from our model; we also regressed each of the 14 outcomes on initial marital status, treatments, and other causal variables used in our analysis of the transition rates. For both our model and the regression model we report R^2, the square of the correlation between the observed and predicted variables. Since we expected a poorer fit from our model than from one designed to minimize errors, we were surprised to find that for 10 of the 14 outcomes our model explains more of the sample variation than does linear regression analysis. Moreover, the advantage of our model in these 10 cases tends to be larger than the advantage of linear regression in the other four.

Though we have considered the predictions of our model for several outcomes at two arbitrary times, we have not yet seen how well it predicts the time path of these outcomes. Computational expense has forced us to examine the time path of only one outcome. We selected the proportion who

TABLE 3

OBSERVED AND PREDICTED VALUES OF ARBITRARY OUTCOMES

Outcome*	Sample† Size	Observed Mean	Our Model: Mean Residual	Our Model: R^2	Linear Regression Model R^2	Difference in R^2
Single at $t = 1$	1,917	351	− 005	.675	664	011
Single at $t = 2$	1,917	.350	002	500	481	019
Married at $t = 1$	1,917	598	011	631	627	.003
Married at $t = 2$.	1,917	540	− 004	439	428	011
Attrited at $t = 1$..	1,917	050	− 006	008	010	−.002
Attrited at $t = 2$.	1,917	111	002	.025	023	.002
No. of marital dissolutions, $t = 1$	1,917	050	− 002	.050	.043	007
No. of marital dissolutions, $t = 2$	1,917	109	−.001	083	068	.015
No. of marriages, $t = 1$	1,917	046	002	111	087	024
No. of marriages, $t = 2$	1,917	086	− 006	138	109	.029
Continuously single to $t = 1$	705	848	004	069	072	− 003
Continuously single to $t = 2$	705	740	.020	092	.089	003
Continuously married to $t = 1$	1,212	885	016	022	.029	− 007
Continuously married to $t = 2$..	1,212	768	017	044	.045	− 001

* Except for the variables on number of marital events, the observed variables are dummy (0–1) variables

† There were 705 initially single white women and 1,212 initially married white women, giving a total of 1,917 white women.

were unmarried at time t, conditional on not having been lost through attrition by time t. We chose this because it is similar to the most important policy outcome; because it should reveal whether experimental effects are confined to an initial, brief adjustment period; and because it depends about equally on our estimates of marriage, dissolution, and attrition rates. This choice provides a severe test of our method because it uses estimates of all four rates of change (r_{12}, r_{13}, r_{21}, r_{23}).

Figure 2 gives observed and predicted curves for this outcome by support level. Points on the observed curve are given by $N_{1j}(t)/[1 - N_{3j}(t)]$ where $N_{1j}(t)$ and $N_{3j}(t)$ are the number in treatment j at time t who are unmarried and lost through attrition, respectively. Points on the predicted curves are calculated as

$$\frac{1}{N_j} \sum_{i=1}^{N_j} \hat{p}_{1ji}(t)/[1 - \hat{p}_{3ji}(t)], \tag{44}$$

where $N_j = N_{1j}(0) + N_{2j}(0)$ (the initial number in treatment j), and $\hat{p}_{1ji}(t)$ and $\hat{p}_{3ji}(t)$ are calculated for each woman i enrolled in treatment j using equations (28) and (30), respectively. Predictions are based on the one-period models of remarriage and attrition and on the two-period model of dissolution; they assume each woman has her assigned treatment.

We begin by considering the observed (squiggly) curves. First, note that the proportion of unmarried women at the start of the experiment differs

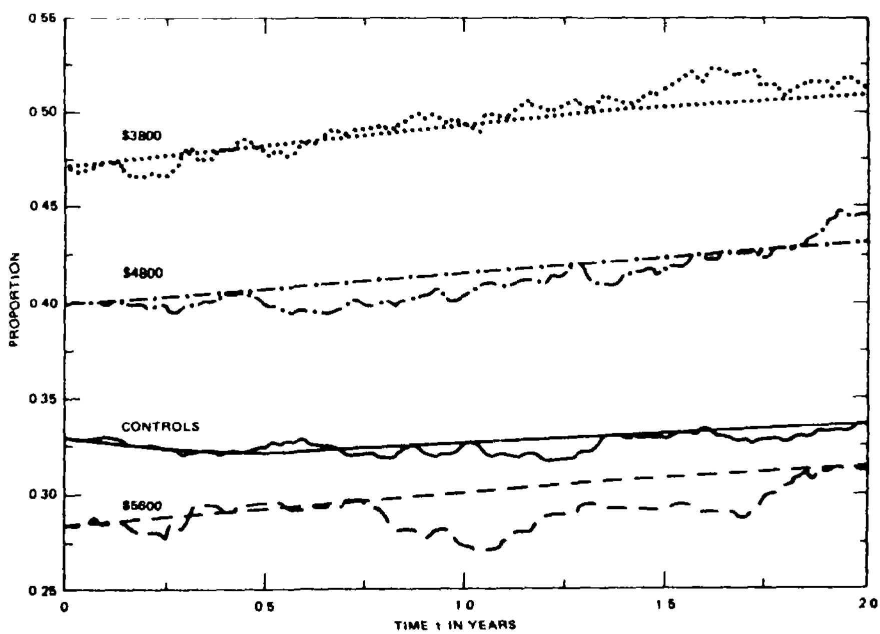


FIG. 2.—Observed and predicted proportion of white women who are single at time t, conditional on not having been lost through attrition by time t, by support level.

greatly from one treatment to another. These initial differences result from the use of a stratified random design in which marital status was a stratification variable. Unmarried women were more likely to be assigned to treatments with a lower support level. Because of this, a comparison of "posttest" levels is clearly inappropriate.

Next, notice that the observed curve for the control group is relatively flat, suggesting that there are no important "natural" time trends (due to aging, secular change, etc.). However, the observed curves for the three support levels show noticeable increases in the conditional proportion of unmarried women after two years (+.039 for the low support, +.044 for the medium support, and +.031 for the high support versus +.009 for the control group). Furthermore, the proportion of unmarried women among those on the financial treatments rises fairly steadily throughout the two-year period. This upward trend is quite apparent for the low and medium supports. It is less certain for the high support, which has the fewest subjects and the most extreme fluctuations about any overall time trend. There is little evidence that the proportion of unmarried women among those on financial treatments has reached a plateau within the first two years of the experiment, as we would expect if an equilibrium was attained during this period. This suggests that dynamic analysis is really needed to assess experimental effects accurately. We elaborate on this in Section III.

Now let us consider the fit between the (smooth) curves predicted by our model and the actual (squiggly) curves. We rely on visual inspection to compare the two sets of curves. On the whole, the fit is quite good except for the high-support group, for which the actual curve is noticeably below the predicted curve. Because only 240 women are in the high-support group, a change in status of a very few women makes a substantial difference in the observed curve. Hence the deviations for this group are less worrisome than they would be for a large group like the control (N = 847).

Our scrutiny of the implications of our model for various outcomes at arbitrary times has revealed no major disadvantages and even some small advantages. The model's primary advantage is, of course, its ability to predict the time path of a variety of interdependent outcomes reasonably well.

III. COMPARISON WITH OTHER APPROACHES

We return finally to an issue raised at the outset: the advantages of event-history analysis relative to other approaches. We continue to assume throughout this discussion that events are generated by a Markov process whose transition rates are log-linear functions of exogenous variables (36).

Obviously any comparison of alternative approaches depends on the assumptions about the process generating events. Following Coleman's (1964, 1968) work, nonetheless, we wish to challenge the still widespread

view that substantive assumptions ought to be dictated by the form in which the data are collected. That is, we do not believe that sociologists ought to change their assumptions about the underlying process (their "model," in our terminology) when they shift from analyzing panel data, say, to analyzing cross-sections or event histories. In our view, a major advantage of formulating problems in terms of dynamic stochastic models is that we can use different data structures to estimate parameters of the same model. This provides a way of unifying a variety of data-analytic procedures.

We begin with an extended discussion of cross-sectional analysis because this has been—and will undoubtedly continue to be—the mainstay of sociological research. We then contrast event-history analysis with two other strategies that use temporal data: event-count analysis and panel analysis. Our discussion of each is brief. To the best of our knowledge, event-count analysis has not yet been developed (let alone applied) except for the most elementary kind of Markov model (a Poisson model). So our comments on it are intended to encourage the development of this approach. However, Singer and Spilerman (1974, 1976) have treated in detail the difficulties of panel analysis with categorical dependent variables. We review these mainly to emphasize that they are either absent or much less serious with event-history analysis.

Cross-sectional Analysis

Cross-sectional data give the state that each member of a sample occupies at a particular time t. Earlier we referred to the unconditional probability of being in a state j at time t as the state probability, $p_j(t)$. Given n possible states, there are only $(n - 1)$ unique state probabilities since the n probabilities must sum to unity. But, in general, there are $n(n - 1)$ unique transition rates. Because $n(n - 1) > (n - 1)$ for $n > 1$, it is immediately obvious that cross-sectional analysis does not allow all parameters of a model to be identified unless we can specify $(n - 1)^2$ of the transition rates, either on theoretical grounds or from a priori knowledge. With event-history analysis we can estimate all parameters. So event-history analysis is clearly preferable to cross-sectional analysis if we wish to understand a process fully or to predict other outcomes.

Under certain circumstances, however, cross-sectional analysis can supply useful information about the process generating events. It is worthwhile to identify these conditions. We begin by considering the situation in which the process has been operating a comparatively long time so that the distribution of the population across states is in equilibrium. For concreteness we again start with the two-state case.

Since the rates, r_{12} and r_{21}, cannot be negative, the two-state model implies that the probabilities of being in each state eventually reach stable, so-

called steady-state, values:

$$p_1(\infty) = \frac{r_{21}}{r_{12} + r_{21}} = \frac{\exp[\theta_{21} x]}{\exp[\theta_{12} x] + \exp[\theta_{21} x]} = 1/\{1 + \exp[(\theta_{12} - \theta_{21}) x]\}, \tag{45}$$

$$p_2(\infty) = \frac{r_{12}}{r_{12} + r_{21}} = \frac{\exp[\theta_{12} x]}{\exp[\theta_{12} x] + \exp[\theta_{21} x]} = 1/\{1 + \exp[(\theta_{21} - \theta_{12}) x]\}. \tag{46}$$

Of course, individuals continue to change from one state to the other. That is, the model implies an equilibrium probability distribution on the aggregate level; it neither assumes nor implies equilibrium on the individual level.

Note that

$$\frac{p_1(\infty)}{p_2(\infty)} = \exp[(\theta_{21} - \theta_{12}) x] \tag{47}$$

or

$$\ln \frac{p_1(\infty)}{p_2(\infty)} = \gamma_{12} x, \tag{48}$$

where $\gamma_{12} = \theta_{21} - \theta_{12}$. For example, this model implies that in the steady state, the log of the odds of being unmarried (rather than married) is linear in x. Equation (48) is the usual form of a binary logit model (Berkson 1944; Theil 1969, 1970), and if all members of x are dummy variables, then it is just a special case of Goodman's (1972) log-linear model of the odds ratio. Thus, when a population is in equilibrium, logit (or log-linear) analysis of cross-sectional data tells us the difference in the effects of variables on the two rates, r_{12} and r_{21}. Note that the finding of "no effect" for a variable in the cross-sectional logit analysis can be due to its equal effects on the two rates. That is, it should not be taken as evidence that a variable is irrevelant to the process, only that it has no net effect on the steady-state distribution.

Is there a similar connection between the general, n-state Markov model and multinomial logit analysis? Unfortunately, the answer is no. This can be proved by a single case. Consider $\ln[p_1(\infty)/p_2(\infty)]$ for the three-state model with attrition, which we have used in our empirical analysis. Though both $p_1(\infty)$ and $p_2(\infty)$ are zero (because eventually everyone "attrites"), they do have a finite ratio. We find that:

$$\lim_{t\to\infty} \ln \left\{\frac{p_1(t)}{p_2(t)}\right\} = \ln \left\{\frac{r_{21} + p_1(0)(r_{23} + \lambda_2)}{r_{12} + [1 - p_1(0)](r_{13} + \lambda_2)}\right\}. \tag{49}$$

The expression on the right-hand side is quite complicated and substitution of equations like (36) does not produce anything resembling a logit model in general. (If $r_{13} = r_{12}$, i.e., attrition rates for married and unmarried women are identical, then [49] does simplify to [48].)

However, an important class of Markov models—general birth and death

models—does have a steady-state distribution that has the form of a multinomial logit model. In these models states can be ordered so that there can occur only transitions between neighboring states (e.g., in a three-state model transitions from one to three and from three to one are impossible). Then, for example, in a three-state model with ordered states, equations (47) and (48) continue to apply, and in addition:

$$\frac{p_2(\infty)}{p_3(\infty)} = \exp\left[(\theta_{32} - \theta_{23})x\right], \tag{50}$$

$$\ln \frac{p_2(\infty)}{p_3(\infty)} = \gamma_{23}x\ , \tag{51}$$

where $\gamma_{23} = \theta_{32} - \theta_{23}$. Furthermore,

$$\frac{p_1(\infty)}{p_3(\infty)} = \frac{r_{32}r_{21}}{r_{12}r_{23}} = \exp\left[(\theta_{32} + \theta_{21} - \theta_{12} - \theta_{23})x\right]. \tag{52}$$

The similarity in form of (52) to (47) and (50) means that there are no clues in cross-sectional data to tell us how to order the states. This must be done on theoretical grounds—or else we must have event-history data, which do permit us to observe what kinds of transitions can occur. But, if we know the order of states in a general birth and death model, cross-sectional multinomial logit analysis does let us make conclusions about the net effect of a variable on transition rates between states.

So far we have considered the situation in which the system is in equilibrium. We have indicated that in the steady state, the log of the odds of being in one state rather than another has a very simple form if states can be ordered. However, if the steady state has not been reached, the log of the odds of being in one state rather than another is a very complicated function of time, the transition rates, and the initial conditions—even for the very simple two-state case. This can be seen by forming the ratio of the right-hand sides of equations (26) and (27).

Suppose we perform cross-sectional analysis at two successive time points and determine that variables have very different effects. We cannot be certain what to conclude. Has the underlying process changed—that is, has the relationship between variables and transition rates altered? Or has the system just moved closer to its ultimate steady-state value? Without some form of temporal analysis we cannot answer these questions.

Social scientists have been so wedded to cross-sectional analysis that they seldom seem to reflect on the likelihood that equilibrium exists or on the length of time required for a system to reach a new equilibrium following some intervention or structural upheaval. We suspect that in most cases inertia greatly slows the speed with which social systems reach equilibrium. We note that the equilibrium assumption implicit in sociological theories prominent a few decades ago (e.g., functionalism) began to be attacked more

than a decade ago. But these criticisms have barely begun to penetrate sociological methodology.

We will make our discussion of this issue more concrete by referring to our empirical illustration. Many analysts faced with data like ours might conduct some sort of logit analysis that assumes that equilibrium is reached within the experimental period. But can we expect the steady-state probability distribution of marital status to be approached during the three or five years of the experimental period? As social experiments go, SIME/DIME is long, so it might seem that this would happen. However, according to the models we have discussed, how long it takes to approach a new steady state (say, to within .01 of its ultimate value) depends on both preexperimental and experimental rates of marital status change. Our results imply that SIME/DIME is much too short for the steady state to be approached during the experimental period.

Marriage and dissolution rates are approximately equal to annual probabilities of forming and dissolving a marriage, respectively.[12] Among SIME/DIME's participants, who have incomes below the U.S. median, both rates are somewhat higher than in the overall U.S. population. In the environment facing the control group, $r_{12} = .10$ and $r_{21} = .05$ are fairly typical.[13] If an aggregate-level equilibrium exists at the beginning of the experiment, the initial probability of being unmarried is $.05/(.10 + .05) = .333$.

Our analyses reported in Section II indicate that for whites SIME/DIME has a negligible effect on marriage rates but roughly doubles the dissolution rate of those on most treatments. If the dissolution rate under income maintenance is twice that of the controls, that is, if $r_{21} = .10$, then according to the two-state model the equilibrium probability of being unmarried under income maintenance is $.10/(.10 + .10) = .50$. Thus the model predicts that under such conditions, the proportion of unmarried women in the population would eventually increase about 50% above its pre-income-maintenance value (from .333 to .500). However, the proportion of unmarried women would increase only by about .04, or about 13%, in the first two years (see fig. 2) and by about .09, or about 26%, in the first five years. It would take nearly 19 years to be within .01 of the steady-state proportion. If data on marital status of participants at any point during the three to five years of the experiment are analyzed cross-sectionally, the ultimate effect of income maintenance will be greatly underestimated. As we mentioned earlier, the

[12] If there is no attrition, the probability of a change in marital status before t is $1 - \exp(-r_{jk}t)$ (see eq. [17]). By a Taylor expansion this is approximately $1 - (1 - r_{jk}t) = r_{jk}t$. So when $t = 1$, the probability of a change is approximately r_{jk}.

[13] These numbers are obtained by rounding off the crude proportion of controls who marry (if initially unmarried) or end their marriages (if initially married) in the first year of the experiment.

observed curves for support levels in figure 2 do not suggest that a plateau or equilibrium has been reached within the first two years.

The only way to decide whether a system is in equilibrium is to collect data over time and to analyze it dynamically—that is, in a way that lets us study the time path of change in the phenomenon.

Event-Count Analysis

A number of sociological studies have analyzed the number of times a particular event occurs in some time period. For example, Spilerman (1970) analyzed the number of riots per city during the mid-sixties. We refer to this type of analysis as event-count analysis. We suspect that a number of surveys have asked such questions as "How many times have you been married?" and "How many times have you been divorced?" However, we are not aware of any event-count analyses in sociology where counts of more than one type of event are analyzed within a single model.

Given the assumptions of a Markov model we can derive expressions for the expected number of different types of events in some time interval. For example, for the two- and three-state models discussed in this paper, we derived the expressions given in equations (31) through (35). These equations, combined with observed data on the counts of different types of events, permit transition rates to be estimated by a nonlinear regression approach. That is, we can estimate rates—or the causal effects of variables on rates—by minimizing the sum over all units of the squared deviation between the observed count of events for each unit and that predicted by the model. This approach, which we have not yet used, has one inherent limitation: We know of no theorem (comparable to the Gauss-Markov theorem in linear regression analysis) that estimators obtained in this way will have optimal statistical properties—even in an infinite sample.

Maximum-likelihood estimators typically have optimal asymptotic properties. But to perform maximum-likelihood estimation we must know the probability-mass function for the count of events. To the best of our knowledge, the expression for this function has not yet been derived for a general n-state Markov model. In fact, it is not even clear that such an expression can be written in closed form for a general n-state model. The probability-mass function for the number of events can, of course, be written explicitly for certain special cases (e.g., a Poisson model), but it is mathematically intractable even for the two- and three-state models used in the illustrations in this paper.

Given these difficulties, it seems obvious that event-history analysis is preferable to event-count analysis. Nonetheless, event-count analysis deserves further study. Under some circumstances, event counts either already exist or are feasible to collect, while event histories or panel data cannot be obtained.

Panel Analysis

Panel data, which record the states occupied by members of a sample at a series of discrete points in time, are the temporal data most commonly available to sociologists. Singer and Spilerman (1974, 1976) have identified the following problems regarding estimation of transition rates in a general n-state Markov model from panel data.[14]

First, sometimes panel data on categorical variables cannot be embedded in—that is, described by—a Markov process. Moreover, sampling variability and measurement error can cause panel data to be unembeddable even though they are truly generated by a Markov process.

Second, even though the panel data may be describable by a Markov process, there may not be a *unique* matrix of transition rates that describe the data. Furthermore, the different matrices obtained in the nonunique cases may suggest substantially different qualitative conclusions (see, for example, Singer and Spilerman 1976, p. 31). Neither embeddability nor uniqueness is a problem in event-history analysis because maximum-likelihood estimators based on such data give unique estimates of rates—or of causal effects on rates.

Third, Singer and Spilerman (1976, pp. 44–48) note that small changes in an observed matrix of transition probabilities (due to sampling variability or measurement error) can sometimes lead to very marked changes in estimates of transition rates. This is clearly undesirable. However, in our experience in analyzing event histories, given a moderately large sample, fairly substantial errors in records on the occurrence or timing of events do not qualitatively alter estimated patterns of causal effects of variables on transition rates.[15] Insensitivity to sampling and measurement error is, we believe, an important advantage of event-history analysis. Because such errors are unavoidable, the sensitivity issue clearly deserves further study—in both panel and event-history analysis.

Fourth, estimation of transition rates from panel data is also sensitive to the length of the time interval between waves of the panel. When the time interval is large, each row of the matrix of transition probabilities approaches the steady-state probability distribution (see, for example, equations [8] through [12]). In this situation there are only $(n - 1)$ unique transition probabilities, rather than $n(n - 1)$. This means that the data contain no more information than cross-sectional data. However, if the time interval

[14] Some of these problems do not arise in certain special cases. For example, the first, second, and fifth problems mentioned below do not occur for the two-state model.

[15] We used this feature in Hannan et al. (1976) to eliminate the possibility that effects of income-maintenance treatment on the rate of marital dissolution were due to attrition bias. Our unpublished Monte Carlo studies show that random error in the timing of events has surprisingly small effects on the quality of maximum-likelihood estimators of rates obtained from event-history analysis.

between waves of the panel is very short, almost all members of the sample will be in their original state. This is not very informative either. With event-history analysis the length of the observation period cannot be too long. It can be too short—if no event has occurred. However, Tuma and Hannan (1978) have shown that with samples that are moderate in size, rates can be estimated well if as few as 10% of the sample have experienced an event. We have not seen comparable results based on panel analysis, but we suspect that it does not perform as well.

Fifth, we know of no way of estimating parameters in a general n-state Markov model from panel data when transition rates are functions of exogenous variables, as is almost always the case in problems of interest to sociologists. Singer and Spilerman (1974) have reported some work on estimating parameters from panel data generated by a mixture of Markov processes, when the mixture is described by some specified probability distribution. This work is helpful, but it still does not permit causal inferences to be made. As we have shown, causal relationships are easily studied with event-history analysis.

A sixth problem with panel analysis concerns the ability to study and detect time dependence in the process generating events. This appears to present a very difficult problem for panel analysis. As we have indicated, various kinds of time dependence can be investigated readily through event-history analysis.

Conclusions

Our conclusions are very simple. Event histories provide rich opportunities for answering fundamental sociological questions. We have shown how to analyze event histories when data are generated by a well-behaved stochastic process. The procedures we have outlined permit analysis of causal effects on the rates at which events occur and of time dependence in such rates. These procedures are simple to implement, and in our empirical application they yield good predictions about a variety of observable variables.

Event-history analysis offers substantial advantages over other common approaches to the study of causal effects on changes in qualitative variables. Since in many situations it is no more difficult to obtain information on the timing of events than the count of events, we urge that sociologists begin to collect and analyze such data.

APPENDIX

To write a likelihood equation like (39) for the model with transition rates that vary from one period p to another (43), we need the survivor function for state j, $G_j(t \mid u, x, z_p)$, where $u = t_{m-1}$ [the time of the $(m - 1)$th event]

and t is some later time, $u \leq t$. The survivor function is obtained by solving the analogue of (21):

$$f_j(t \mid u, x, z_p) = r_{jp} G_j(t \mid u, x, z_p), \tag{A1}$$

where $r_{jp} = \Sigma_k r_{jkp}$. Since $f_j(t \mid u, x, z_p) = -dG_j(t \mid u, x, z_p)/dt$, this implies that

$$\frac{dG_j(t \mid u, x, z_p)}{G_j(t \mid u, x, z_p)} = -r_{jp} dt. \tag{A2}$$

Integrating both sides from u to t, we obtain

$$\ln G_j(t \mid u, x, z_p) = -\int_u^t r_{jp} dt \tag{A3}$$

$$G_j(t \mid u, x, z_p) = \exp\left[-\int_u^t r_{jp} dt\right]. \tag{A4}$$

To eliminate the integral on the right-hand side of the expression above, define u'_p and t'_p as follows:

$u'_p = 0$ if $\tau_p < u$ or $t < \tau_{p-1}$ (state j is entered after period p ends or t occurs before period p begins);

$= u$ if $\tau_{p-1} \leq u < \tau_p$ (state j is entered in period p) ;

$= \tau_{p-1}$ if $u < \tau_{p-1} < t$ (state j is entered before period p begins and t occurs after period p begins);

$t'_p = 0$ if $\tau_p < u$ or $t \leq \tau_{p-1}$ (state j is entered after period p ends or t occurs before period p begins);

$= \tau_p$ if $u < \tau_p < t$ (state j is entered before period p ends and t occurs after period p ends);

$= t$ if $\tau_{p-1} < t \leq \tau_p$ (t occurs in period p);

where τ_p is the end of period p, as defined in the text. Using these definitions,

$$G_j(t \mid u, x, z_p) = \exp\left[-\sum_{p=1}^{P} (t'_p - u'_p) r_{jp}(x, z_p)\right]. \tag{A5}$$

So, analogous to (39), we can write the likelihood equation for the model given by (43) as:

$$\mathcal{L} = \prod_{j=1}^{n} \prod_{i=1}^{N} \prod_{m=1}^{\infty} \left\{[G_j(t_m \mid t_{m-1}, x, z_p)]^{v_{m-1,j}} \cdot \prod_{k=1}^{n} [r_{jkp}(x, z_p)]^{w_m v_{m-1,j} v_{mk}}\right\}. \tag{A6}$$

REFERENCES

Albert, A. 1962. "Estimating the Infinitesimal Generator of a Continuous Time, Finite State Markov Process." *Annals of Mathematical Statistics* 33 (June): 727–53.

Bartholomew, David J. 1957. "A Problem in Life Testing." *Journal of the American Statistical Association* 52 (September): 350–55.

———. 1973. *Stochastic Models for Social Processes* 2d ed. New York: Wiley.

Berkson, J. 1944. "Application of the Logistic Function to Bioassay" *Journal of the American Statistical Association* 39 (September): 357–65.

Blumen, Isadore, Marvin Kogan, and Philip J McCarthy. 1955. "The Industrial Mobility of Labor As a Probability Process." Cornell Studies in Industrial and Labor Relations, no. 6. Ithaca, N.Y.: Cornell University Press.

Breiman, L. 1969. *Probability and Stochastic Processes*. Boston Houghton Mifflin.

Coleman, James S. 1964. *Introduction to Mathematical Sociology*. Glencoe, Ill : Free Press.

———. 1968. "The Mathematical Study of Change." Pp. 428–78 in *Methodology in Social Research*, edited by H. Blalock and A Blalock New York McGraw-Hill.

———. 1973. *The Mathematics of Collective Action*. Chicago. Aldine.

Cox, D. R., and H D. Miller 1965 *The Theory of Stochastic Processes* London. Methuen

Dhrymes, Phoebus. 1970 *Econometrics: Statistical Foundations and Application* New York: Harper & Row.

Feller, William 1968 *An Introduction to Probability Theory and Its Applications*. Vol 1. 3d ed. New York: Wiley.

Goodman, Leo A. 1972. "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables" *American Sociological Review* 37 (February): 28–45

Hannan, Michael T, Nancy Brandon Tuma, and Lyle P Groeneveld 1976. "The Impact of Income Maintenance on the Making and Breaking of Marital Unions: Interim Report." Research Memorandum 28 Menlo Park, Calif Stanford Research Institute, Center for the Study of Welfare Policy.

———. 1977. "Income and Marital Events: Evidence from an Income-Maintenance Experiment." *American Journal of Sociology* 82 (May). 1186–1211.

Karlin, Samuel, and Howard Taylor. 1975 *A First Course in Stochastic Processes* New York: Academic Press.

Land, Kenneth C. 1971. "Some Exhaustible Poisson Process Models of Divorce by Marriage Cohort." *Journal of Mathematical Sociology* 1 (2): 213–32

McFarland, David D 1970 "Intragenerational Social Mobility as a Markov Process. Including a Time-Stationary Markovian Model That Explains Observed Declines in Mobility Rates." *American Sociological Review* 35 (June): 463–76

Mayer, Thomas F. 1972. "Models of Intragenerational Mobility." Pp 308–57 in *Sociological Theories in Progress*, edited by J Berger, M Zelditch, and B Anderson Vol 2 Boston. Houghton Mifflin.

Singer, Burton, and Seymour Spilerman 1974 "Social Mobility Models for Heterogeneous Populations" Pp. 256–401 in *Sociological Methodology 1973-1974*, edited by Herbert L. Costner. San Francisco: Jossey-Bass

———. 1976. "The Representation of Social Processes by Markov Models" *American Journal of Sociology* 82 (July): 1–54

Sørensen, Aage B. 1975. "The Structure of Intragenerational Mobility" *American Sociological Review* 40 (August): 456–71

Spilerman, Seymour. 1970. "The Causes of Racial Disturbances: A Comparison of Alternative Explanations." *American Sociological Review* 35 (August): 627–49.

———. 1972*a*. "The Analysis of Mobility Processes by the Introduction of Independent Variables into a Markov Chain" *American Sociological Review* 37 (June): 277–94.

———. 1972*b*. "Extensions of the Mover-Stayer Model." *American Journal of Sociology* 78 (November): 599–626.

Theil, H. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10 (October): 251–59

———. 1970. "On the Estimation of Relationships Involving Qualitative Variables." *American Journal of Sociology* 76 (July): 103–54.

Tuma, Nancy Brandon. 1976 "Rewards, Resources, and the Rate of Mobility: A Non-Stationary Multivariate Stochastic Model." *American Sociological Review* 41 (April): 338–60.

Tuma, Nancy Brandon, Lyle P. Groeneveld, and Michael T. Hannan 1976. "First Dissolutions and Marriages: Impacts in 24 Months of SIME/DIME." Research Memorandum 35. Menlo Park, Calif.: Stanford Research Institute, Center for the Study of Welfare Policy.

Tuma, Nancy Brandon, and Michael T. Hannan 1978, in press. "Approaches to the Censoring Problem in Analysis of Event Histories." In *Sociological Methodology 1979*, edited by Karl Schuessler. San Francisco: Jossey-Bass.

Tuma, Nancy Brandon, Michael T Hannan, and Lyle P. Groeneveld 1977 "Variation over Time in the Impact of the Seattle and Denver Income Maintenance Experiments on the Making and Breaking of Marriages " Research Memorandum 43. Menlo Park, Calif.: Stanford Research Institute, Center for the Study of Welfare Po.icy.

Normative Expectations and the Emergence of Meaning as Solutions to Problems: Convergence of Structural and Interactionist Views[1]

Warren Handel
Southern Illinois University at Edwardsville

Structural and interactionist approaches to role theory are often considered fundamentally incompatible. This paper argues that the two are compatible and complementary. Structural role theory has come to regard inter- and intrarole conflict as the natural state of affairs; thus conformity to internalized norms has been undermined as an explanatory device and as the image of the actor's generic orientation to the normative order. In structural theory. systemic stability and patterned conduct are explained by structural mechanisms that ameliorate the adverse consequences of conflicting expectations. The question of how the actor copes with conflicting expectations is not addressed in middle-range structural theory, although it is implied as a conceptual problem. Interactionists have addressed the negotiation of meaning in interaction as the actor's practical solution to the problems caused by conflicting social pressures. Negotiated meanings do not replace conflicting expectations, but coexist with them as a working consensus among actors concerning how conflicts are to be resolved in particular situations, despite their several preferences. The development of the working consensus and the symbolic representation of conflicting models of what ought to be the case are implications of the model of the pragmatic actor developed by George Herbert Mead.

It is commonly held that the structural and interactionist approaches to role theory are essentially or paradigmatically different.[2] Interactionist thought

[1] This paper would not have been possible without the substantive and editorial help of my colleague Judith Handel at every stage of its development. Sabbatical leave from Southern Illinois University at Edwardsville made the writing more comfortable than it might have been. The paper was much improved by the thoughtful and constructive comments of two anonymous referees and the editors.

[2] The term "interactionist," unless otherwise qualified. refers to the symbolic interactionist position. The shorter, less precise term is used in deference to some authors cited, whose studies are conceptually compatible with symbolic interaction but who would probably be uncomfortable to be designated as symbolic interactionists The term "structural" is used to refer to the family of role-theoretic models that has evolved from the work of Talcott Parsons and Ralph Linton Merton's work will receive the most specific attention, but the major arguments apply to any of the variants that recognize intra- and interrole conflict as the natural state of affairs It is important to emphasize, following Merton (1957, 1967), Goode (1973*a*), and Gans (1972) that, in addition, there is no fundamental conflict between structural and conflict-theoretical sociology as they have been practiced recently.

is characterized, especially from outside the specialty, as psychologistic, astructural, and sufficiently indeterministic to preclude, in principle, a deductive, predictive social science. Structural thought is criticized for placing such an untenable explanatory burden on conformity to norms internalized during socialization that, in principle, it underrecognizes the degree of social conflict and cannot account adequately for social change. Often, structuralists are held to ignore the special characteristics of humans and human society in a self-conscious attempt to emulate the natural sciences.[3] Only occasionally, and without supporting argument, has it been asserted that the structuralist and interactionist variants of role theory are compatible (Goode 1973*b*; Merton 1967, p. 43) and complementary (Merton 1975). This paper is an attempt to document the latter view by countering the usual characterizations of structural and interactionist role theories.

Inter- and intrarole conflict have come to be generally recognized in structural theory as the natural state of affairs. One theoretical consequence, most explicitly developed by Merton (Merton 1957; Merton and Barber 1963), has been the displacement of conformity to internalized norms as an explanatory device and even as the image or model of the actor's generic orientation to the normative order. The phenomenon to be explained, in fact, has been reformulated as the maintenance of order despite conflicting normative expectations.

Conflicting normative expectations pose a continuing problem for the actor in assessing the consequences of possible courses of action and coordinating the intended course of action with those of others. The problem of how actors cope with the conflict, while implied by structural theory and recognized as an important topic for study, falls without the domain of middle-range structural role theory.

Interactionist role theory has been concerned with the negotiation of emergent meanings in interaction as part of the process of coordinating actors' anticipated courses of action. The emergent definitions are conceived as solutions to practical problems arising, for the most part, from patterned, conflicting social pressures. The existence of such pressures is an important part of the environment in which the negotiating occurs, but this structural environment is seldom studied systematically by interactionists as a topic in its own right. Conceived as a response to conflicting social pressures, the

[3] These characterizations have become ritualized and are not considered to require careful statement or support Two illustrations exemplify the ritualization. Stokes and Hewitt (1976) address their discussion of aligning actions to a remedy of exactly these theoretical flaws and to a reconciliation of the structural and interactionist approaches. However, in an earlier paper (Hewitt and Stokes 1975, p 10) they conceded that their own formulation of structural theory might be considered a caricature. Stinchcombe (1975, p. 14) attributes the "hopeless proposition, 'God only knows what definition people will give to situations,'" to "much contemporary work" in the tradition derived from Mead and Thomas, without further specification.

negotiation of emergent meanings serves as an image of actors' generic orientation to the normative order that is compatible with recent structural role theory. In addition, negotiation among actors is a conceptually suitable answer to the question of how actors cope with conflicting normative expectations.

Thus, the structural and interactionist approaches to role theory are well, although not uniquely, suited to one another. The major phenomena addressed in each approach are compatible with the models of the other. And, insofar as one approach directly addresses the patterning of expectations and the other the orientation of actors to those expectations as a practical problem, each addresses problems implied, but not addressed, by the other.

NORMATIVE CONSISTENCY AND THE EXPLANATORY POWER OF INTERNALIZED NORMS IN STRUCTURAL ROLE THEORY

Consistent Normative Expectations as the Cornerstone of Normative Explanation

The explanatory power attributed to internalized norms follows from both strategic (or rational) and moral considerations. Insofar as ego and alter conform to the norms associated with their respective positions, each will be performing in accordance with the expectations of the other, and, therefore, each will favorably sanction, and be favorably sanctioned by, the other. Further strategic benefits arise indirectly from the smooth functioning of the social system implied in conformity to its mandated division of labor. In addition to these strategic concerns, the internalization of norms as moral or value prescriptions leads to some degree of observance of them for its own sake. The confluence of moral and strategic concerns accounts for the conformity of each actor to the norms and provides for systemic stability through the self-perpetuating character of mutual conformity and mutual reward.

The explanatory strength of this argument, however, rests on the assumption that the rationally and morally concerned actor is confronted with situations in which the applicable norms are consistent. If the norms are inconsistent, so too will be the flow of sanctions. Actors may continue to orient themselves strategically to the distribution of sanctions, but the consequence of systemic stability will no longer follow. Further, the ability of actors to engage in rational or strategic planning in a social world of inconsistent sanctions will itself need explanation instead of serving to explain systemic stability.

Similarly, each actor may continue to conform to norms because they are held to be morally correct. However, if the actors are diverse in their moral principles, the effects of moral adherence to selected norms will not be sys-

temic stability. Instead, actors will find themselves in moral disputes. Further, the moral and strategic considerations will contradict rather than reinforce one another to the extent that each actor's preferences are sanctioned in accord with others' contrary values.

The Explanatory Task Posed by Sociological Ambivalence

Conformity to norms internalized through socialization has been regarded historically as the generic orientation of actors to the normative order and as the basic explanatory principle of structural role theory with respect to both the conduct of actors and systemic stability. But in a recent, systematic review of Merton's theory of social structure, Stinchcombe (1975, p. 12) has characterized the core process in Merton's view of social structure as "the choice between socially structured alternatives."[4] Merton (1975, pp. 34–35) asserts, further, that social structures are characterized essentially, and as a stipulation of his approach to structural analysis, by "incompatible patterned expectations and a 'dynamic alternation of norms and counter-norms.' "[5] Insofar as the normative structure is essentially composed of pairs of opposing norms,[6] any action that has meaning in the community can be found to be simultaneously supported and condemned, prescribed and proscribed, by appropriately selected legitimate norms.

Internalization of an ambivalently structured normative system cannot explain the actor's conduct. A norm is not simply a formulation of a course of action. It is, rather, a formulation of a preferred course of action for a specific class of actors in a specific class of situations. Thus, an ambivalent social structure is not one that alternately values contrary courses of action, but one that simultaneously values contrary courses of action for a single

[4] Kelly (1974, p 288) concludes from his study of kinship and residence rules among the Etoro that the "organization of contradiction is the essence of social structure." In Etoro society he found conflicting expectations that were resolved by a variety of devices in making particular decisions. His analysis is especially important because it extends the domain of a model of social structure based on conflicting expectations to primitive society and because it was developed relatively independently of the similar models in sociology.

[5] Cancian (1975) differentiated norms into three types: ranking norms that serve as standards for evaluating performance; reality assumptions that define the domain of culturally meaningful or possible actions; and membership norms that set standards for the inclusion of a person in a social group or position. An illustrative ranking norm and counternorm pair is "good fences make good neighbors" and "something there is that doesn't love a wall"; a reality assumption and counterassumption pair is "birds of a feather flock together" and "opposites attract"; and an opposing pair of membership norms is "the public is welcome" and "we reserve the right to refuse service to anyone."

[6] This should not be taken to exclude partial conflict or other relationships among norms For example, the norm "when in Rome, do as the Romans do" is only sometimes in conflict with the norm "a day's work for a day's pay." Each has a counternorm as well, "to thine own self be true" and "whatever the traffic will bear" or *"caveat emptor,"* respectively.

actor in a given situation. Consider this opposing pair of military strategies: "Discretion is the better part of valor" and "Damn the torpedoes! Full speed ahead." If the latter is understood to apply only in situations in which one must turn broadside before tail and in which it is better to bring the enemy within range of one's guns sooner than late, the qualification of situation removes the apparent contradiction between the formulations of conduct. Similarly, the apparent contradiction is removed if the former is understood to apply only to indispensable personnel or the latter is understood to apply only to personnel on certain assignments, such as mine-field sweeping or bomb disposal. To constitute a pair of opposing norms, the behavioral prescriptions must remain contrary when all such boundary conditions of their application are considered. Without denying the importance of norms, or of the sanctions associated with them, some explanation must be provided for the actors' selection of particular norms as appropriate. That explanation cannot resort to conformity to internalized norms and must, therefore, be nonnormative in the sense of not following the normative pattern of explanation (Wilson 1970).

Nor can internalization of an ambivalently structured system of norms explain systemic stability. Without mechanisms to articulate their choices, an aggregate of actors, freely and separately choosing between contrary norms, would be rent by internal moral conflict. Despite the operation of articulating mechanisms, actual systems are ubiquitously characterized by moral dispute and even, occasionally, by civil war. Without them, the conflict would escalate in intensity and complexity to the point at which order could not be maintained. These conflicts, grounded in moral commitment, are not closely akin to the Hobbesian war of each amoral, rational actor against all the others. Shifting coalitions of zealots and moral entrepreneurs would prevail, limited in issues under dispute only by the cultural supply of alternatives and in treachery only by the degree to which agreement on some moral issues leads actors to temper the enmity arising from their disagreement on others. Of course, the degree to which such compromises are proper would be another, and pivotal, moral question within each coalition of actors agreeing on other matters, but not all. Fittingly, as the underlying dynamic of social order shifts from conformity to a set of consensual values to the articulation of choices among conflicting ones, the nature of disorder shifts as well, from the amoral, rational pursuit of short-run hedonic goals to zealous moral crusading.

Finally, insofar as the actor cannot conform to the normative structure without simultaneously violating it, conformity to internalized norms has become an inappropriate image of the actor's generic orientation to the normative structure. In fact, since any act whatsoever can be characterized accurately as either the violation of a norm or conformity to its counter-norm, the designation that conformity or deviance has occurred signifies

nothing about the act itself. The choice of designation, however, may reflect the preferences of the designator and the relationship between designator and actor and may be socially consequential for all concerned.

The changes in structural role theory have been sufficiently drastic to be characterized as a change in the "ideal of the natural order" (Toulmin 1961). The theoretical transformation followed from increasing attention to role conflict until inter- and intrarole conflict became accepted as the natural state of affairs. The continuing need for solutions to the explanatory problems thus posed has been explicitly noted as an implication of empirical results and theoretical considerations.

Haga, Gruen, and Dansereau (1974) used longitudinal data to study the development of role conceptions among a cohort promoted to new positions. The data did not fit a model based on conformity to fixed roles but suggested instead that the credibility of that model might be an artifact of nonlongitudinal study designs. Room (1976) reviewed the attempts to explain patterns of alcoholism in terms of sociological ambivalence. He concluded that without detailed specification of the mechanisms involved in its resolution, sociological ambivalence has little explanatory power and raises more questions than it resolves. Coser (1975) argued that in complex, ambivalent structures the actor is afforded more autonomous choice than in simple ones and will more frequently develop flexibility and self-direction as personal characteristics. "Complex role-sets and differentiated roles are . . . [the] basic structural precondition of individuality" (Coser 1975, p. 259). The attribution of autonomy to actors, however, does not explain their conduct, but acknowledges our ignorance of its causes.

Explaining Conduct Subject to Interrole Conflict

Initially, studies of role conflict did not question the internal consistency of the normative expectations, or role, associated with a single position in the social structure. Conflicting expectations were recognized, however, as an implication of the simultaneous applicability of normative expectations associated with the multiple positions held by an actor in a single situation. By the early 1950s interrole conflict had emerged as a major empirical topic and theoretical issue for structural role theorists. Later, it became the most common single topic in studies of role structure.

In response to early studies of role conflict, Toby (1952) argued that interrole conflict is the normal state of affairs and that every society must have institutionalized means to reduce its disruptive effects for individuals and for the social system. Some of the institutionalized means he proposed, especially the excuse, subsumed conflicting expectations under a broader consensual hierarchy of values that ordered the expectations and resolved the conflict. These mechanisms assumed that the choice between conflicting

expectations was socially organized by conformity to a set of more general consensual norms. These mechanisms, then, continued to explain order by conformity to internalized norms.

However, at the same time, Toby (1952) proposed mechanisms that did not rely on normative consensus to resolve role conflicts. He argued, for example, that lying could mitigate the effects of role conflict and might sometimes be valued as tact. In addition, he suggested that the actor might take a creative role in resolving conflicting demands by, for example, stalling, playing off one group against another, or effecting a redefinition of the conflicting roles.

The early studies of interrole conflict, then, resulted in differentiation of the explanatory power of roles taken singly from that of roles taken collectively. Each role was regarded as internally coherent and, accordingly, as an explanation of conduct through the mechanism of conformity. Taken collectively, however, roles were no longer regarded as coherent and were seen to present problems for the actor that required strategic efforts without the normative framework. Manipulative skill was suggested as an important element of behavior associated with roles.

Explaining Conduct Subject to Intrarole Conflict

The characterizations of roles considered singly and collectively were redintegrated in Merton's (1957) discussion of the role set: he extended the principle that conflicting expectations are the natural state of affairs to the conception of the internal structure of single roles. The role set represents the application of Marxian and Durkheimian principles to the study of social microstructures. Merton asserts that the Marxian principle that "social differentiation generates distinct interests among those variously located in the structure of society" (Merton 1957, p. 112) is fundamental to all sociological theory. Applied to the role set as a social microstructure, this principle implies that the occupants of the various other-roles associated with any given position should exhibit systematic differences in their expectations concerning the occupants of that position. Following Durkheim, Merton concerns himself with the articulation of these diverse expectations in the system—with the rate at which and the manner in which expectations associated with various positions are honored. The overall result, especially when later extensions (Merton and Barber 1963) are included, is a model of social structure that is thoroughly based on the articulation of conflicting normative expectations, as opposed to conformity to consensual ones.

The articulating mechanisms proposed by Merton (1957) suggest a social factual pattern of explanation. Orderly conduct, conceived as predictable patterned response to conflicting social expectations, is to be explained by other social factual conditions that operate to reduce the variance in what

expectations are honored. For example, there is a distribution of power among actors. Insofar as that distribution is patterned by position, occupants of those positions that have greater power will be more likely, by virtue of that occupancy, to have their expectations honored. To be sure, there are normative expectations concerning who is entitled to power—and counternorms as well. It is the actual distribution of power that articulates the role set, however. Similarly, the other mechanisms proposed by Merton can be resolved to the distribution of other variables—for example, observability and solidarity. The articulation of the role set is the combined effect of these variables on the pattern of responses to conflicting normative expectations.

Merton (1957, p. 118) seeks to exclude "the utopian figment of a perfectly effective social system." The recognition of sociological ambivalence (Merton and Barber 1963) extends the principle of normative conflict to include conflict among the expectations expressed by the several occupants of a single status. This implies that no actor can precisely anticipate the expectations of another on structural grounds. Structurally informed anticipation of what any other actor expects can be no more precise than the statistical distribution of expectations allows. And that varies from one system and from one norm/counternorm pair to another.

Because of this ambivalence, too, the articulating mechanisms cannot, in principle, precisely determine a common course of action for the several actors in a status. The articulating mechanisms adapt conduct to real environmental conditions. Adaptive conduct, however formulated, involves an evaluative element. The ambivalence of a class of actors implies that, even if their common position exposes them to common experiences, they will evaluate those experiences differently and, therefore, respond differently.

Three evaluative variables are especially important. First, the actors will differ in their own preferences. Second, they will differ in their evaluations of others' responses to their conduct. Third, they will differ in their evaluation of the importance of adhering to one's own preferences in the face of negatively valued responses. Applied simultaneously to occupants of all statuses, these variables imply that no actor can anticipate precisely, by virtue of his structural circumstances, what will be expected of him by others, by virtue of their structural circumstances. Further, even conceding that the same "best guess" anticipation of others' responses is reached by a group of actors, because of their own value differences they will not act alike. This reciprocally reduces certainty based on structural information.

Merton (1957) self-consciously limits his discussion to the structural articulation of conflicting expectations in order to illustrate the middle-range theory strategy. The question of "how individuals happen to deal with the complex structures of relations in which they find themselves" (Merton 1957, p. 112) is explicitly excluded from consideration. However, the recog-

nition of residual conflict as the natural state of affairs implies this question as a crucial one for a more general theory. Despite the operation of structural articulating mechanisms, an ambivalent social structure must confront the actor as a chronic problem. How the actor addresses that problem and determines a course of action, then, is a critical theoretical issue.

Although its solution cannot be structural, it is theoretically important to characterize the problem posed for the actor by an ambivalently structured system of expectations. The structural indeterminateness implied by sociological ambivalence does not imply that information of a nonstructural sort cannot greatly reduce, or even eliminate, the uncertainty. For example, biographical information about particular others in their uniqueness, about social types (Klapp 1958), or about the particular setting could be used. People are not chronically immobilized by indecision despite the uncertainty in our structural models. Characterizations of how they deal with the conflict extend the domain of our theories and may improve their precision.

The actor is likely to have internalized ambivalent expectations and to experience moral doubt. In estimating the response of others strategically, the actor will be confronted with contradictory expectations and anticipated responses from more or less powerful, more or less informed, more or less concerned others. Morally and strategically, a course of action will not be ambiguously defined by the normative structure and the sanctions associated with it. Instead, the normative structure will confront the actor with both risk and uncertainty to the degree that the conflicting expectations are not systemically articulated. Insofar as residual conflict remains, one's problem is to assess the situation rationally and to improve one's fate in it by the application of manipulative skills to the problem of mitigating the adverse consequences of choosing one sanctioned course of action over another. The actor will be interested in knowing which norms, of those supportable in a situation, will actually be supported in each instance.

STRUCTURAL EMBEDDEDNESS OF INTERACTIONIST STUDIES

Recognition of Structural Constraints in Diverse Theoretical Formulations

Interactionists have been concerned with specifying the social processes through which meaning arises in social interaction. The emphasis has been on flexibility and change in social arrangements, rather than on their stability. In addition, although social interaction is recognized as occurring in broader social environments, theoretical attention has been focused primarily on the interaction process itself, relatively exclusive of those environments. Finally, although interaction processes are conceived to produce meaning, attention has been largely confined to formulating the processes, with little attention paid to specifying systematically the meaning that arises.

Despite relative consensus concerning the topic of study, however, there is relatively little explicit theoretical agreement among interactionists. Especially, no particular body of work or single statement is collectively acknowledged to represent the interactionist position, and there is relatively little terminological consistency. Therefore interactionists are frequently characterized as sharing a perspective or orientation rather than a theoretical position. Herbert Blumer is often taken as representative of the interactionist community, but accurately so only with respect to his influential programmatic and critical statements (see Blumer 1954, 1956, 1962, 1966). Far from claiming that his work is theoretical in any strict sense, Blumer has consistently held theory as a programmatic goal and taken those holding other theoretical perspectives to task precisely for their premature theoretical pretensions. (For a sense of the diversity in theoretical formulation among interactionists, cf. Meltzer, Petras, and Reynolds [1975], Morris [1977], and Swanson [1968]. For a relatively integrative statement, see Lauer and Handel [1977], especially chaps. 1–3.)

For the present argument, the crucial matter is the way in which the broader social environment is regarded by interactionists—even though a common feature of their perspective is a relative lack of theoretical interest in that environment. Relative theoretical disregard of a topic need not be confused with a denial of its existence or importance, although it often is. The diversity of interactionist formulations has not precluded a remarkable degree of consensus on one matter. Each characterization of the interaction processes recognizes that the structural context in which interaction occurs constrains, even determines, those processes.

This recognition has its roots in the seminal work of G. H. Mead and W. I. Thomas. Mead's (1964*d*, 1964*e*) programmatic evaluation of the social sciences and of the place of social psychology was based on his belief that laws of social organization would eventually be formulated that would be equivalent to physical laws. Social consciousness was conceived to presuppose a determining social order. The task of social psychology was to articulate objective social conditions, stated in terms of social laws, with the social consciousness they produce. Both in the evolution of our species and in the lives of its members, the social relationship has logical and temporal priority over the development of corresponding selves and reflective minds.

The Thomas dictum (Thomas and Thomas 1928, p. 572) is sometimes understood to imply a nondeterministic position. It is embedded, however, in a discussion of the situational method that embraced the behaviorism of Watson and Pavlov as foundational for sociology. Behaviorism was applauded for providing an explanation of the organism's behavior in terms of its situation. The term "situation" refers to the total circumstances of the organism—immediate matters and distal ones alike, for example, the concrete responses of others and the parents' educational levels.

Sociology was viewed as the special case of situational explanation in which the impersonal stimuli originate in the social situation. Thomas (1927) regarded the social situation as overwhelmingly determinant of action and even of creative thought (1909), but he left the question of complete determinism open. As situations become more complex, he argued (Thomas 1927; Thomas and Thomas 1928), statistical analysis becomes essential to uncovering the causes of conduct in them. In a secular application of Pascal's wager, he advocated statistical analysis coupled with qualitative case studies to generate hypotheses for further, more refined, statistical analyses. This program is appropriate regardless of the ultimate ability of the statistical analysis fully to explain conduct.[7]

Macrostructural constraint of negotiated meaning.—The recognition of structural constraints continues to be an explicit feature of otherwise diverse formulations. A few important, and terminologically diverse, statements will serve as illustrations. A convenient starting point is papers by Shibutani (1970) and Farberman (1975) that go beyond merely acknowledging the effects of structural variation and explicitly incorporate macrostructural conditions in the core of their arguments. Farberman (1975) argues that the market conditions of the automobile industry, including the pricing of new cars, the costliness of maintaining inventories, interest rates, the distribution of knowledge among consumers, the sales-tax structure, and attitudes toward that structure, combine to generate criminal practices in used-car sales and a set of meanings to describe and justify them. The complex arrangements among the participants in the used-car business clearly involve negotiated meanings, but the important constraints on the outcomes of the negotiations are macrostructural market conditions that lead to similar practices and meanings in the several separate dealerships studied.

Shibutani (1970) argues that intersocietal conflict is reflected in the norms governing interaction and that, as a result, the meanings developed in interaction are constrained in a way that aids in the maintenance of morale necessary to continue the conflict. Specifically, he argues that in situations of prolonged intersocietal conflict the enemies' acts are systematically interpreted in the worst possible light, while those of one's own side are interpreted as favorably as possible. Sources that deviate from these biases are denounced as traitors or dupes, leading contrary statements to be disregarded, if made, and discouraging people from making them. As a

[7] Pascal's wager is a pragmatic proof that one ought to believe in God. Pascal conceded that the question of God's existence was not amenable to proof or certain resolution. He argued, though, that if God does not exist no harm would be done by erroneous belief in His existence If God does exist, however, erroneous disbelief may have drastic consequences Thus, one ought to believe in God Pascal's wager is the prototype of the pragmatic approach to insoluble questions The approach is to evaluate the consequences of alternative premises under varying truth conditions and to accept the least onerous risk. Thomas's proposed methodology exhibits pragmatic consideration of the consequences of acting on belief in free will and determinism.

result, the opponent group is increasingly characterized as evil, while one's own side is increasingly characterized as good—ultimately justifying a contrast conception in which the opponent group is only marginally human, embodying all the traits held in disfavor by one's own. By extending the argument to include, although generally in milder form, various intrasocietal conflicts such as those between racial or ethnic groups, Shibutani clearly indicates that the causal agent is not the patriotic fervor of war, but the structurally based bias implied in intergroup conflict. Patriotic fervor, racism, and other antagonistic meaning structures are the result of structural arrangements, not their cause.

Structural constraint of interaction process.—Turner's (1962) discussion of role making is often regarded as the exemplar of the interactionists' astructural approach. It is concerned with the process through which one's own and imputed other-roles are created in interaction as interlocking gestalts with which to interpret and guide conduct. Turner, however, does not argue that this process is a replacement for models based on relatively routinized following of preset roles. Rather, he asserts that the location of situations on a continuum of degrees of institutionalization affects the flexibility accorded participants and the very nature of their involvement in role conduct. The rigid conformity that characterizes the military and the flexibility of role making are equally particular to ranges of situations differentiated by their respective degrees of structurally imposed rigidity and equally dictated by them. Thus, although his discussion is concerned in detail only with role making, the process is recognized to be consequential in only certain kinds of situations and to be only one value of a continuum of ways in which actors can be involved in role behavior, each way particular to, and constrained by, the structural context of interaction.

Scheff's (1968) discussion of negotiating reality is concerned with the way in which institutionally based power can be employed to control the outcomes of negotiating meanings in interaction. The degree to which power is asymmetrically distributed among interactants, however, is recognized as a variable; thereby the techniques employed and the degrees of control in interaction are related to the social positions of participants. Goffman's (1959) classic study of self-presentation explicitly relates expressive conduct in interaction to organizational goals and constraints, broad cultural values, and the availability of organizational resources. Shibutani (1966) relates both the accessibility of others to conversation and the range of acceptable content to the collective need for news. The need for news is a reflection of the degree to which existing organizational structures are functioning effectively. His earlier work (Shibutani 1961) recognizes that even the degree to which one can express personality characteristics is structurally constrained.

Thus, while the formulations of interaction processes are diverse, they

recognize in common that those processes, however characterized, are constrained by the broader social environment. That environment is defined with respect to such standard sociological concerns as distribution of power, effectiveness of existing structural arrangements, rigidity and specificity of organizational norms, market structures, degree of conflict. While the formulations are quite diverse, the tendency of interactionists to concern themselves in any given statement with interaction in a narrow range of situations may tend to exaggerate the conceptual differences among them. However, that assessment must await restatement of the various formulations in a neutral terminology and careful examination of whether contrary characterizations of processes are made of interaction in similar settings.

Structure in the Logic of Interactionist Studies

Both theoretically and methodologically, interactionist studies take into account the broader social structural contexts that are the core subject matter of sociology. Three ways in which interaction is theoretically recognized as being embedded in broader social structural contexts are especially important. First, as a premise for negotiating particular obligations, standards of performance, use of available resources, etc., more or less standardized cultural expectations are theoretically asserted. These serve as the actors' definition of a situation in advance of firsthand knowledge about the specific situation and its participants. And, as situated information becomes available, revised local knowledge is employed in that setting until contradicted. Thus, whether the knowledge is part of the broad cultural background or situated, interaction processes have existing contextual knowledge as a premise from which they proceed. Second, actors are generally conceived to be accountable to perform roles, whose details may require negotiation, but which provide incumbents with typical goals, resources, and obligations These roles constitute the world of the actor within which negotiations may be conducted and under whose auspices interaction occurs Third, the typical matters of concern in a type of interaction—forms of expression, consequences of the interaction, typical sequences and connections among interactions, types of individuals involved and their fates, among other matters—are all constrained by the broader social context in which interaction occurs. These structural considerations are not necessarily discussed systematically or extensively in interactionist theory, but neither are they denied or discounted. Rather, they are recognized to pertain and alluded to as the context or environment of the interaction processes that are addressed in detail.

The recognition of broader social structural context is reflected methodologically as well. In the most general sense, the social situation, methodologically presumed to be an ongoing reality, gives substantive meaning to con-

duct occurring within it and makes it possible for that conduct to be recognized as an instance of particular social processes. First, particular acts derive their meaning, at least in part, from the situation in which they occur. Even those acts that serve to redefine the meaning of events and to make meaning fluid in a situation are understood to mean what they do only with reference to the context in which they occur. The removal of tires from an automobile has one meaning in a service station, another on a dark city street. Similarly, acts are recognized analytically as instances of particular processes by reference to their context. For example, the reading of newspapers by musicians is recognizable as an instance of backstage behavior if it occurs in the orchestra pit of a theater out of the audience's view during a performance (Goffman 1959). Reading a newspaper in a college classroom expresses role distance, but not in a backstage region; it simultaneously establishes identity characteristics, provides a miniature backstage region for the concealment of facial expression and the direction of gaze, and evaluates the instructor. Similarly, the fact that a singles dance is a low-prestige occasion makes it possible to recognize conversations there as exchanges of sad tales and face-saving (Berk 1977).

Normative Expectations, Working Consensus, and Standard Operating Procedures in Ethnographic Analysis

Social arrangements and definitions are regarded in interactionist studies as emergent products of social interaction in a setting. However, a relatively more stable system of expectations is generally recognized as an important factor in constraining those emergent arrangements. Typically, ethnographic studies are organized around a single substantive theme that pervades activities in a setting or around the creative response to the contingencies imposed on a given position in the system. Occasionally, too, a small system such as a court will be abstracted from its social context and studied as an acting unit. Attention is focused on the practical problems confronting the people in the setting. Their ability to sustain a theme, meaning, or concerted activity and their ability to develop integrated role and career lines are analyzed as solutions to these problems. Nearly all the practical problems can be formulated as conflicting demands or definitions of the situation imposed on the actors by others in the situation. That is, the emergent definitions and standard operating procedures are solutions to the problems implied for the actor by social structure characterized by intra- and inter-role conflict.

Very significantly, the emergence of collective solutions does not imply extensive revision or removal of the conflicting demands or definitions. These continue to be pressed and thus to stabilize the emergent meaning and roles as a situated routine or standard operating procedure. Accordingly,

new participants in the situation will, through socialization or their own problem-solving activities, arrive at an accommodation with these routines, their own preferences notwithstanding. New personnel may, of course, press slightly different claims from those already suggested, leading to modifications of the routinized resolution of the conflicts inherent in the situation. In fact, most interactionist studies do not describe the actual emergence of the working consensus and standard operating procedures. Instead, these are presented as stages in the natural history of existing arrangements as they are currently encountered and appear to have developed.

Rather than by the revision or removal of conflicting demands, conflicts are resolved by the development of a working consensus (Goffman 1959, pp. 9–10) to which participants adhere without strong commitment. This working consensus, an emergent collective product, exists as an organizing principle in and for the group without necessarily displacing the particular preferences of individuals or being expressed by any one of them (Shibutani 1966). Even for those who would prefer the situation to be handled otherwise, the working consensus constitutes knowledge of the way things are done in the situation and, therefore, the way they must be done to avoid disturbing the routine.

In the interactionist studies, then, two distinct modes of social organization are implied. Competing normative claims are pressed and constrain activities in a manner compatible with Merton's concept of the articulation of role sets. In addition, a working consensus and standard operating procedures are negotiated that stabilize or routinize activities in the setting. In effect, the negotiated meanings reduce the discretion afforded participants by structural arrangements. This reduces uncertainty and risk concerning the way in which activities in the particular setting will be conducted. Participants need not subscribe normatively to these negotiated settlements. Practical problems in the setting, of which the normative competition is the most sociologically interesting, sustain them as solutions.

Illustrations of the Logic of Interactionist Analyses

Especially in attempting to characterize a broad and terminologically diverse literature, a "for instance" is not a proof. Still, this formulation of the logic of interactionist studies will be supported by two examples. Their authors, Erving Goffman and Ralph Turner, are often regarded as exemplifying polar extremes within the interactionist community.

Because they interpret Turner's (1962) study of role making as a general formulation of interaction processes, rather than as one applying to a limited domain, and because they ignore additional contrary implications of his other work, other sociologists frequently characterize him as an especially thoroughgoing astructural or radical interactionist. In contrast, if the

interactionist position is characterized as radically astructural, a selective reading of Goffman's work, which is explicitly attentive to structural matters, can lead to characterizing him as a structuralist and not at all an interactionist (cf. Gonos 1977). Thus, the convergence of their views on one aspect of structure can be regarded as a strong indication of agreement on that issue among interactionists on the continuum between them. Such a convergence can be seen in the importance of structural considerations in the two articles summarized below.

Turner's (1947) study of disbursing officers.—Although serving in the military, the disbursing officer is personally accountable and financially liable for expenditures in deviation from regulations, even if that deviation is directly ordered by a superior officer. The problems facing the disbursing officer arise from his peculiar social circumstances: he is often given orders that conflict with regulations; his clients frequently outrank him; and there is a network of informal structures that exerts pressure on him because of his formal position and duties. According to Turner the career and operating style of the disbursing officer are solutions to the problems posed by these conflicting pressures. Four career patterns emerge—each representing a different routinized solution to common problems and each accomplished by defining a distinctive type of role relationship with the others with whom he works.

Turner's (1947) discussion contains all the elements that still characterize interactionist ethnographies. First, people are located in a social circumstance in which the persistent pressures imposed by others, formally and informally, confront them as a problem or as a sequence of consequential choices. The choices are not made once and for all, but must be continually affirmed as a career line, since the pressures continue to be present and demands continue to be pressed. The career takes the form of developing a routinized, mutually satisfactory style of relationship with the others in one's environment. Especially important to the thesis of this paper is Turner's finding that the three types of solution that do not deal realistically with the totality of pressures involved are infrequent and unsuccessful. His study thus implies that the structural situation has consequentiality on its own terms and limits the range of successful solutions.

*Goffman's (1961*a*) study of institutional underlife.*—Underlife develops in total institutions as a response to their peculiar organizational characteristics, especially the contrast between the identity implications of institutional arrangements and those preferred by inmates. Mere participation in the institution acknowledges its definition of one's identity to a degree, but the underlife expressively denies it, even when the material benefits involved are not large.

The underlife is not a catalog of discrete, idiosyncratic or symptomatic rule evasions, but a cooperative exchange system. Extra haircuts, showers,

laundry privileges, and tailoring are exchanged among inmates for each other or for small sums of money, cigarettes, or candy. The services are materially important in themselves, and the recognition by others that the recipient is entitled to them is crucial to the expression of identity. This is especially clear in such territorial claims as appropriating a comfortable chair. The chair is good to sit on, but the important identity implications arise from the uncontested, if grudging, recognition of the territory.

The staff is involved in the exchange as well. By tacit means, such as ignoring or tolerating violations, and active ones, such as hiring inmates or exchanging friendly and public greetings with them, they establish extrainstitutional relationships with inmates. This implies ratification of identity claims by a less suspect source. In exchange, they achieve greater cooperation in the maintenance of social control. For example, inmates refrain from "acting up" when guests are on the ward.

The underlife, then, represents the collective solution of institutionally generated problems—inmate needs for ratification of their identities and staff needs for cooperation in maintaining discipline. The material arrangements may take varied forms, but the acknowledgment of extrainstitutional identities is routinized at a level and in manners constrained by both the institutional arrangements and the persistent demands of the inmates.

Clearly, in the two studies just considered, Turner's view of structure is compatible with Goffman's, and neither view is incompatible with that of the structuralists.

Normative Expectations in Identity Negotiations

Role distance (Goffman 1961*b*) is an increasingly important concept in the study of identity expression. It presumes that mere incumbency in a social position, in the absence of expressive conduct to the contrary, has standardized identity implications for the incumbent. Much of the negotiation of identity of self characteristics consists of the expression of distance from those standardized identity implications and of preferred alternatives.

The standardized identity implications of incumbency in a position may constitute a complete characterization of the incumbent (Goffman 1961*a*), an important but limited aspect of a single organizationally situated identity (Emerson and Pollner 1976), or a minor and relatively transitory attribution (Berk 1977; Davis 1975). However, in each case, the identity implications are found waiting as expectations held for any incumbent in a position by virtue of that incumbency. In addition to standardized identities, the situation provides the resources with which negotiations are conducted and places constraints on the interactions in which the identities are implied and negotiations conducted.

The distribution of identities is a critical element of the working consen-

sus and not a matter left to individual preference. Since the identity of each participant in the situation has identity implications, and other consequences, for the other participants, all will be involved in selectively ratifying the claims of others and attempting to impose a distribution of identities favorable to their own plans (Weinstein 1969). The result is an integrated system of identities, consisting of both organizationally given and distance elements, which are ratified in and organize the activities of situated activity systems. Each identity includes what the person claims, and is expected, to be in the situation by its participants, including the duties, official and unofficial, liked and disliked, that he performs routinely and accountably. In effect, then, the system of situated identities is a way of formulating the division of labor in a particular setting; it defines the standard operating procedures in a situated activity system.

The core of Goffman's (1961*b*) discussion is his attempt to contend with the dual theoretical observation that role distance is "not part of the normative framework of role" (1961*b*, p. 115) but is nonetheless "almost as much subject to role analysis as are the core tasks of roles themselves" (1961*b*, p. 152). Recognition of the situated activity system as a locus of social organization serves to show how minute social organization can be shaped and maintained without corresponding minute normative expectations. In their stead, actors orient to concrete, situated, pragmatic concerns.

Coser (1966) recognized the structural significance of role distance and its compatibility with Merton and Barber's (1963) later discussion of sociological ambivalence. However, she attempted to assimilate the concept without modifying the structural approach to accommodate the situated activity system (Goffman 1961*b*) as a locus of nonnormative social organization additional to the normative structure. In so doing, she redefined the concept, somewhat reducing its range of applicability and all but eliminating the interactional components. Thus, her recognition that different strains of sociological theory could be fruitfully combined was flawed by a failure to acknowledge and accommodate their substantive differences.

Three definitional changes imposed by Coser on the term "role distance" are especially noteworthy. First, by linking the concept to anticipatory socialization she (Coser 1966) imposes restrictions on the taking of role distance. As Stebbins (1967) observed, she thereby excludes role distance within stable role sets, as exemplified by Goffman's discussion of surgeons. Activities that Goffman identified as role distance were reanalyzed as normative. Also eliminated is nostalgic role distance, which asserts claims to roles no longer held, whether they were more complex and demanding than current ones (as in retirement, demotion, or status degradation) or more simple (as when role distance represents the failure to assume new responsibilities).

Second, Coser's (1966) discussion limits the concept to the affectation of role behavior associated normatively with one's actually anticipated roles.

This eliminates role distance that asserts false claims, either cynically or innocently. For example, the antics of lower-middle-class girls while horseback riding are role distance, for Coser, because they can conceive of themselves as possibly attaining the higher status they burlesque. If we assert, rather, that they conceived that they could pass as upper class, not attain that status, or that they simply took pleasure from burlesquing what they were not and knew they would not be, or that we were ignorant of their thoughts, their conduct would no longer be role distance, as defined by Coser (1966).

Finally, in direct contradiction of Goffman, Coser (1966) asserts that role distance is normative. This tends to undermine the possibility of addressing the situated activity system as a locus of organization distinct from the normative structure. The issue here is not the relative utility of the two approaches to role distance, only the difference between them.

The Interactionist Program and the Domain of Sociological Theory

Turner (1962, p. 38) described the interactionists' programmatic task as one of formulating the "social psychological mechanisms through which the functional principle of social stability operates." By addressing the negotiation of meaning by actors and their development of satisfactory, if not gratifying, standard operating procedures in response to problems posed by patterned conflicts in the pressures or expectations placed on them, interactionists have progressed toward that goal. Negotiated definitions and reasonably stable standard operating procedures constitute, for the actors, knowledge concerning what will be done by others in the immediate circumstances, which of their duties will be scrutinized and sanctioned, and what the boundaries of their discretion are. This situated knowledge increases the accuracy of strategic calculations by reducing uncertainty concerning the consequences of anticipated courses of action. Further, despite the continued pressing of normative preferences by participants, the working consensus and standard operating procedures are sufficiently stable and consensual to organize a routine to which mutual orientation will be mutually satisfactory.

Confronted with a structure of expectations to which conformity is impossible, actors negotiate a situated order, habitual conformity to which will produce satisfactory results, and continually renegotiate elements of that order as the need arises—but only as the need arises. This understanding of the interactionist approach does not relegate it to the subsidiary task of "interpreting" structural findings. Instead, it extends the domain of sociological theorizing to include a sociological account of how actors accommodate themselves to the structural indeterminance with which they are confronted, even when structural articulating mechanisms are taken into

account. Interaction is, intrinsically, a collective endeavor and negotiating is, intrinsically, a social activity.

IMPLICATION OF STRUCTURE IN THE MODEL OF THE PRAGMATIC ACTOR

The model of social structure characterized by conflicting expectations implies, as a theoretical problem, the consideration of how the actor addresses the practical problem posed by the conflicting expectations and sanctions. As a matter of fact, the major focus of interactionist research has been the manner in which emergent definitions and standard operating procedures serve as practical solutions to essentially that problem. The interactionists are concerned with that problem as a matter of principle as well, because, as I shall now argue, it is implied in the model of the pragmatic actor as it appears in the work of George Herbert Mead.

The pragmatic actor is morally and strategically concerned and is given, whenever possible, to habitual conduct. Habitual conduct is characterized by the repetition of behavior patterns that have proved satisfactory in prior application to similar circumstances and by relative inattentiveness to minor variations in the situation and to possible alternative behavior patterns.

The conduct of the pragmatic actor is cybernetically controlled by the continuous comparison of events, especially the results of ongoing conduct, with models of what ought to be the case. These models serve as goals for purposeful conduct; distance from those goals serves as the impulse to modify one's conduct (Mead 1964*a*). The actor is interrupted in this habitual attitude and becomes conscious of the situation only insofar as it disappoints expectations. Imaginings, on such occasions, consist of projected courses of action and their consequences. These imaginings allow the actor to project conditions to what they will have become after an anticipated action and so to evaluate prospectively its effectiveness in reducing the distance between models of what ought to be and the perceived reality. When the existing situation approximates the model, the actor behaves habitually, continuing ongoing conduct without thought or consciousness and employing patterns of conduct that have proved successful in similar situations. When disappointment goads the imagination, conduct that is prospectively successful is tested in the world and, insofar as it remains successful in retrospect, adopted as new habit (Mead 1938).

Actors are free to imagine or define objects in any manner that pleases them. But an object is an anticipated course of action and the consequences of acting on those definitions, while real, need not be the expected or preferred ones. Mead (1964*b*) refers to an accurate anticipation of the consequences of conduct as the measure of the truth of the definition of objects. The world of objects is constructed for action in a social world with ob-

durate characteristics that allow many acts, but not all imaginable acts, to be successfully completed.

The existence of more than one actor of the pragmatic type has interesting sociological implications. No matter how freely each actor arrives at his or her own models of what ought to be, each actor's model is, insofar as it guides action, a condition imposed on the action of the others. Successful planning in the social world requires that each actor orient himself or herself to the actions of others, and insofar as those actions are guided by models, the models themselves are conditions to be considered. Habitual action in the social world requires that one's own actions do not sufficiently disturb the habitual actions of the others that they are led to alter their habits in a way that disturbs one's own. One mechanism, held as a distant goal, that could accomplish the sustenance of reciprocally satisfactory habits is the development of shared models of what ought to be the case and the consequent control of conduct by shared goals (Mead 1964*a*).

The existence of more than one pragmatic actor has this additional implication: even if actors are individually free to define objects as they please, the freedom of each actor is a condition of the action of the others. In redefining any object, and thereby changing one's goals, one reconstitutes oneself as an object predictable to others. Each actor must be able to accommodate plans of action to changes in the plans of others, by altering the principles organizing his or her own conduct in response to a change in the environment—redefining the other.

This redefinition of others as objects can be accomplished retrospectively, by examining a failure of one's own plans that seems caused by an incorrect understanding of those of the others. However, to redefine others prospectively, thus avoiding errors and their consequences, requires an additional mechanism. That mechanism must allow each actor to anticipate the actions of others, even if those actions are being guided by newly developed models that have not yet been revealed in habit. The gesture is the mechanism proposed by Mead to fulfill this social function. It is a bit of observable conduct that indicates anticipated courses of action. Gestures in the conduct of others indicate their plans, even if they are new, untried plans. Reciprocally, others anticipate one's own conduct by observing gestures.

A special subset of gestures, significant symbols, allows each actor to anticipate what plans will be attributed to him or her by others as a result of his symbolizing and hence to communicate purposefully. A collection of symbolizing actors can indicate their plans to one another, know more or less what they have communicated, and reciprocally modify the expressed plans until a mutually satisfactory plan is achieved. With luck, this can be done before the action unfolds to the point that the differences among the plans put the actors at cross-purposes. The preferences of some actors may remain unacceptable to others with whom they must contend. If the former

settle for a plan that runs counter to their own preferences, the group will have arrived at a working consensus.

A system of symbols must meet several criteria to serve this function. First, the system must be sufficiently stable and consensual that uncontested portions of it can be employed by actors to contest and align other portions of it, that is, to introduce new symbols or change the definitions of existing ones as the need arises. Second, insofar as the symbols are not genetically fixed they must be learned through interaction in the community that employs them. Third, the articulation of actual events with symbols must be sufficiently fluid that actors can remain mutually intelligible while defining given events as objects of different kinds. That is, with respect to issues of what is and what ought to be, the symbol system must allow for expression of contrary views, of norms and counternorms, as well as means to reconcile them. In fact, the expression of alternate positions in dispute can be considered the symbol's main function.

The existence of more than one pragmatic actor, then, has several sociologically interesting implications. First, for such actors, "the intelligible order of the world implies a determined moral order—and for a moral order we may substitute a social order, for morality has to do with the relations of intelligent beings with each other—and . . . this determined moral or social order is a world as it should be and will be" (Mead 1964*c*, pp. 248–49). Second, intelligible communication of this world of moral expectations must be possible, even if actors differ about the nature of that world and the nature of their actual circumstances. Third, pragmatic actors will utilize symbolic representations of the moral order to arrive at a collective and more or less mutually satisfactory representation of what should be the case and what, therefore, should be done and by whom. Finally, being habitual, actors will tend to honor existing arrangements, rather than to seek alternatives, as long as they remain satisfactory. That is, the pragmatic actors will tend to construct and honor standard operating procedures as a set of interlocking habits in response to differences with other actors concerning what ought to be the case.

CONCLUSION

The interactionist and structural approaches to role theory are neither what they have commonly been characterized to be nor very different from one another. Recent models of social structure incorporate inter- and intrarole conflict (sociological ambivalence) as the natural state of affairs. Accordingly, an actor cannot conform to all of the normative expectations simultaneously pressed by others. Structural mechanisms articulate the conflicting expectations systemically, but the articulation is imperfect and the actor is confronted with the practical problem of coping with the residual con-

flicts. Residual conflict is intrinsic to the structural model and constitutes a boundary of its domain. It is, in fact, a specific instance of the more general formal principle that probabilistic models are indeterminate with respect to each instance of their application, however exact they may be with respect to rates of occurrence in the population of instances.

Interactionists have empirically addressed the negotiation of situated meanings and standard procedures in interaction as the actors' collective solution to the practical problem of conflicting social pressures. In so doing, they have extended the domain of sociological theory to include routinized social conduct that is sustained, and reorganized when necessary, despite the structural indeterminacy of specific actions in specific circumstances.

A system of symbolic representations of opposing models of what ought to be (norms and counternorms) and the use of those symbolic representations to make possible the development of satisfactorily interlocking habits are implications of Mead's model of the pragmatic actor. The two approaches are compatible and complementary.

A more precise statement of the relationship between these two approaches can be hazarded. The structural approach analyzes the distributions of normative preferences and the resources for their enforcement by structural position. The resolution of conflict is addressed structurally as well. Interactionists have addressed the distributions of preferences and resources by persons in particular social settings and the resolution of their conflicts by negotiations that are constrained by broader structural arrangements. The latter topic is unavailable to the structural approach and thus constitutes an expansion of the domain of sociological theory, but not in the sense of providing a more general unified theory. For, as I have argued, the negotiations become available as a topic only by the presumption of a structural environment. A more general theory would need to incorporate both modes of analysis in a unified conceptual framework or, less likely, replace them.

The theoretical conception of the normative structure is in need of serious reconsideration. The norm is generally conceived as a prescription for the conduct of a type of actor in a type of situation. Our theoretical discussions ought, at least, to cope with paired contrary prescriptions and the implied continuum between them, as its base unit, and might profit from considering collections of normative prescriptions in more complex relationships with one another than simple opposition.

Although normative prescriptions cannot bear the explanatory burden often imposed on them, they do appear to serve important functions. Social norms and counternorms serve as legitimate, if conflicting, formulations of preferences for the purposes of negotiating what arrangements will prevail in a given instance. (For an especially clear discussion of this function of normative prescriptions, see Bilmes [1976].) Second, people may refer ellip-

tically to their estimate of prevailing expectations, the most strategic course of action, standard operating procedures, etc., in normative terms. Statements such as "you should dress up for that party" may serve as variably accurate summaries of complex and contradictory contingencies—that is, the bottom line. Our complex models of conflicting expectations, structural articulating mechanisms, and uncertainty-reducing negotiations can explain, among other things, how such elliptical remarks can make the sense they do (and as much of it). Finally, the expression of preferences or the formulation of conduct in terms of normative expectations may bring mechanisms of accountability to bear. For example, those responsible for enforcing norms may be recruited as third parties to an arrangement; the chance that they will be recruited may be invoked as a bargaining device; or the actor's commitment to the system of norms may be affirmed when it appears doubtful.

Mulkay (1971) arrived at a rather gloomy appraisal of middle-range strategy with respect to its production of deductive theory and methodological rigor. The convergence of two perspectives, long held to be fundamentally different, suggests another dimension for evaluating that strategy—the degree to which, in general, the separate groups of specialists have been converging on formulations that will ultimately prove more general and powerful than their component parts. That such convergence will occur is a fundamental assumption of the middle-range strategy, without which it makes little sense to suspend general theorizing and still hope for a coherent sociology. Especially in the light of disappointing progress in other respects, the degree of convergence in the discipline is a crucial dimension for its evaluation, but cannot be assessed unless synthetic theorizing becomes more common relative to analytic.

This attempt to synthesize structural and interactionist role theories was occasioned, in large part, by the frequency and vigor of the assertions that they are incompatible. While each specialty was changing substantially, and converging on a common model to boot, the characterizations of each by practitioners of the other remained static and misinformed during 2 decades or more of mutual polemic criticism. The intellectual costs of the polemic have been high. Verbal confusions hamper communication between the two camps. Interactionists' failure to explicitly consider structure has contributed to a lack of consistent terminology and to great difficulty in comparing the diverse interaction processes they address. Although the norm remains a fundamental unit of structural thought, it has been stripped of much of its explanatory power, but structural theorizing has been uninformed of the functions of norms observed by interactionists.

Sociology is characterized by many disputes between specialties, and it is unlikely that no others are based on misapprehensions similar to those discussed above. The general point at issue beyond the specific case is that,

if we continue to ground our positive assertions in critical analyses of others' views, we ought to be as careful in formulating our colleagues' views as our own.

REFERENCES

Berk, Bernard. 1977 "Face-Saving at the Singles Dance" *Social Problems* 24 (June): 530–44.

Bilmes, Jack 1976 "Rules and Rhetoric Negotiating the Social Structure in a Thai Village." *Journal of Anthropological Research* 32 (Spring) 44–57.

Blumer, Herbert 1954 "What Is Wrong with Social Theory?" *American Sociological Review* 19 (February): 3–10

——— 1956 "Sociological Analysis and the Variable." *American Sociological Review* 21 (December): 683–90.

——— 1962. "Society as Symbolic Interaction" Pp. 179–92 in *Human Behavior and Social Processes*, edited by Arnold Rose Boston Houghton Mifflin

———. 1966 "Sociological Implications of the Thought of George Herbert Mead." *American Journal of Sociology* 71 (March): 535–44

Cancian, Francesca 1975 *What Are Norms?* London Cambridge University Press.

Coser, Rose. 1966 "Role Distance, Sociological Ambivalence, and Transitional Status Systems" *American Journal of Sociology* 72 (September) 173–87

——— 1975 "The Complexity of Roles as a Seedbed of Individual Autonomy." Pp. 237–63 in *The Idea of Social Structure· Papers in Honor of Robert K Merton*, edited by Lewis Coser New York· Harcourt Brace Jovanovich

Davis, Fred 1975 "The Cabdriver and His Fare Facets of a Fleeting Relationship" Pp. 142–50 in *Life as Theater*, edited by Dennis Brissett and Charles Edgley. Chicago: Aldine.

Emerson, Robert, and Melvin Pollner. 1976 "Dirty Work Designations: Their Features and Consequences in a Psychiatric Setting" *Social Problems* 23 (February)· 243–54

Farberman, Harvey 1975 "A Criminogenic Market Structure. The Automobile Industry" *Sociological Quarterly* 16 (Autumn) 438–57.

Gans, Herbert 1972 "The Positive Functions of Poverty" *American Journal of Sociology* 78 (September)· 275–89.

Goffman, Erving 1959 *The Presentation of Self in Everyday Life* Garden City, N Y. Doubleday

——— 1961*a*. *Asylums* Garden City, N Y · Doubleday

———. 1961*b*. *Encounters*. Indianapolis: Bobbs-Merrill.

Gonos, George 1977. "'Situation' versus 'Frame' The 'Interactionist' and 'Structuralist' Analyses of Everyday Life" *American Sociological Review* 42 (December): 854–67.

Goode, William 1973*a*. "Functionalism: The Empty Castle." Pp. 64–94 in *Explorations in Social Theory*, edited by William Goode. New York. Oxford University Press.

———. 1973*b* "The Social Topography of the Field." Pp 3–32 in *Explorations in Social Theory*, edited by William Goode. New York Oxford University Press.

Haga, William, George Gruen, and Fred Dansereau 1974 "Professionalism and Role Making in a Service Organization: A Longitudinal Investigation." *American Sociological Review* 39 (February): 122–33.

Hewitt, John, and Randall Stokes 1975. "Disclaimers" *American Sociological Review* 40 (February): 1–11.

Kelly, Raymond. 1974. *Etoro Social Structure*. Ann Arbor· University of Michigan Press.

Klapp, Orrin. 1958 "Social Types: Processes and Structure." *American Sociological Review* 23 (6): 674–78.

Lauer, Robert, and Warren Handel. 1977 *Social Psychology: The Theory and Application of Symbolic Interactionism*. Boston: Houghton Mifflin.

Mead, George Herbert. 1938. *The Philosophy of the Act*. Chicago: University of Chicago Press.

———. 1964*a*. "The Genesis of the Self and Social Control." Pp. 267–93 in *Selected Writings,* edited by Andrew Reck. Indianapolis: Bobbs-Merrill.
———. 1964*b*. "A Pragmatic Theory of Truth." Pp. 320–44 in *Selected Writings,* edited by Andrew Reck. Indianapolis: Bobbs-Merrill.
———. 1964*c*. "Scientific Method and the Moral Sciences." Pp. 245–66 in *Selected Writings,* edited by Andrew Reck. Indianapolis: Bobbs-Merrill.
———. 1964*d* "Social Psychology as Counterpart to Physiological Psychology." Pp. 94–104 in *Selected Writings,* edited by Andrew Reck. Indianapolis: Bobbs-Merrill.
———. 1964*e*. "What Social Objects Must Psychology Presuppose?" Pp. 105–13 in *Selected Writings,* edited by Andrew Reck. Indianapolis: Bobbs-Merrill.
Meltzer, B., J. Petras, and L. Reynolds. 1975. *Symbolic Interactionism: Genesis, Varieties, and Criticism.* Boston: Routledge & Kegan Paul
Merton, Robert. 1957. "The Role-Set· Problems in Sociological Theory." *British Journal of Sociology* 8 (June): 106–20.
———. 1967. "On Sociological Theories of the Middle Range." Pp. 39–72 in *On Theoretical Sociology,* edited by Robert Merton New York: Free Press.
——— 1975 "Structural Analysis in Sociology." Pp 21–52 in *Approaches to the Study of Social Structure,* edited by Peter Blau. New York: Free Press.
Merton, Robert, and Elinor Barber 1963. "Sociological Ambivalence." Pp. 91–120 in *Sociological Theory, Values, and Sociological Change: Essays in Honor of Pitirim A. Sorokin,* edited by Edward A. Tiryakian. New York: Free Press.
Morris, Monica. 1977. *An Excursion into Creative Sociology.* New York: Columbia University Press.
Mulkay, Michael Joseph. 1971. *Functionalism, Exchange, and Theoretical Strategy.* New York: Schocken.
Room, Robin 1976 "Ambivalence as a Sociological Explanation: The Case of Cultural Explanations of Alcohol Problems." *American Sociological Review* 41 (December): 1047–65.
Scheff, Thomas 1968. "Negotiating Reality: Notes on Power in the Assessment of Responsibility." *Social Problems* 16 (Summer): 3–17.
Shibutani, Tamotsu 1961. *Society and Personality.* Englewood Cliffs, N.J : Prentice-Hall.
———. 1966 *Improvised News: A Sociological Study of Rumor* Indianapolis: Bobbs-Merrill.
———. 1970. "On the Personification of Adversaries." Pp. 223–33 in *Human Nature and Collective Behavior: Papers in Honor of Herbert Blumer,* edited by Tamotsu Shibutani. New Brunswick, N.J.: Transaction
Stebbins, Robert. 1967. "A Note on the Concept of Role Distance." *American Journal of Sociology* 73 (September): 247–50.
Stinchcombe, Arthur 1975. "Merton's Theory of Social Structure." Pp. 11–33 in *The Idea of Social Structure: Papers in Honor of Robert K. Merton,* edited by Lewis Coser. New York· Harcourt Brace Jovanovich.
Stokes, Randall, and John Hewitt. 1976. "Aligning Actions." *American Sociological Review* 41 (October): 838–49.
Swanson, Guy 1968. "Symbolic Interaction." Pp 441–45 in *International Encyclopedia of the Social Sciences,* edited by David L. Sills. Vol. 7 New York: Macmillan.
Thomas, W. I. 1909. *Sourcebook for Social Origins.* Chicago: University of Chicago Press.
———. 1927. "The Behavior Pattern and the Situation." *American Sociological Association Publications* 22: 1–13.
Thomas, W I., and D. S. Thomas. 1928. *The Child in America.* New York: Knopf.
Toby, Jackson 1952. "Some Variables in Role Conflict Analysis." *Social Forces* 30: 323–27.
Toulmin, Stephen. 1961. *Foresight and Understanding: An Enquiry into the Aims of Science.* New York: Harper & Row
Turner, Ralph. 1947. "The Navy Disbursing Officer as a Bureaucrat." *American Sociological Review* 42 (June): 342–48.

———. 1962. "Role-Taking: Process vs. Conformity." Pp 20–40 in *Human Behavior and Social Processes,* edited by Arnold Rose Boston· Houghton Mifflin.

Weinstein, Eugene. 1969. "The Development of Interpersonal Competence." Pp. 753–75 in *Handbook of Socialization Theory and Research,* edited by David Goslin. Chicago: Rand McNally.

Wilson, Thomas. 1970. "Conceptions of Interaction and Forms of Sociological Explanation." *American Sociological Review* 35 (August): 697–710.

Capital Punishment, Gun Ownership, and Homicide[1]

Gary Kleck
Florida State University

Several issues are examined: (1) the deterrent (or other preventive) effect of the death penalty, (2) the relationship between the level of gun ownership and the homicide rate, and (3) the incapacitative effect of imprisonment on the homicide rate. Findings consistently show a significant reciprocal relationship between gun ownership levels and the homicide rate, with crime pushing up gun ownership and gun ownership in turn pushing up the homicide rate. Greater certainty of arrest and conviction for homicide has a significant negative effect on the homicide rate, suggesting a deterrent effect of legal sanctions, while the homicide rate in turn has a negative effect on the certainty of punishment. This is consistent with a hypothesis of the strain on criminal justice system resources resulting from increasing crime. Further, the hypothesis that the imprisonment of criminal offenders has an incapacitative effect on homicide rates independent of deterrence effects is supported. Contrary to the recent findings of Ehrlich (1975*b*), the analysis fails to detect consistent evidence of a deterrent effect of capital punishment on U.S. homicide rates for 1947–73. This may be because of the infrequency of executions during this period.

This paper attempts to answer two controversial questions, both related to the problem of interpersonal violence in America: (1) Does use of the death penalty exert any measurable influence on the rate of homicide in the U.S.? and (2) What relationship, if any, exists between the level of gun ownership and the level of homicidal violence?

THE DEATH PENALTY

While the subject of the death penalty seems to be of great public interest, some academic critics have argued that it is not an *important* issue. Norval Morris and Gordon Hawkins, in a recent book on crime control, seemed to be disappointed that the public was focusing so much attention on what they dismissed as an issue of "singular inconsequence" (1977, p. 82). Their evaluation seemed to be based on the facts that the death

[1] For his encouragement, advice, and generosity with his time, I wish to thank David J. Bordua. I would also like to thank Alan Lizotte and Kenneth Land, who read previous drafts and contributed many helpful suggestions.

penalty has never affected more than a small minority of convicted murderers and that it has been used infrequently. However, the principal utilitarian justification for the existence of capital punishment has been that it provides a unique deterrent to potential killers. In this light, one would think that the practical significance of the death penalty as a crime control measure would be evaluated in terms of the estimated number of killings prevented because of its use, not the proportion of convicted killers executed or the frequency of executions. Isaac Ehrlich's recent methodologically sophisticated analysis of the death penalty's deterrent effect suggested that the estimated average trade-off between executions of an offender and the lives of potential victims over the period 1933–69 was of the order of magnitude of one for eight (1975*b*, p. 414). If one assumes that Ehrlich is correct, how significant would the resumption of use of the death penalty be in reducing homicide? As recently as 1947 there were 4.26 executions per 100 estimated convictions for murder and nonnegligent manslaughter. For this risk of execution to have pertained in 1974, when there were an estimated 21,465 homicides and an estimated 6,980 convictions for murder and nonnegligent manslaughter, there would have had to be 297 executions. If eight homicides had been prevented for each execution, then this hypothetical execution risk would have resulted in the saving of about 2,376 lives, that is, about 11% of the homicide victims in 1974. These hypothetical figures are presented for purely illustrative purposes in order to demonstrate the concrete implications of certain aspects of the death penalty debate. Assuming Morris and Hawkins would not consider a policy which could conceivably save over 2,300 lives a year an issue of "singular inconsequence," their evaluation seems to have been based on an unspoken assumption that capital punishment does not deter homicide to any significant degree. It is precisely this assumption which needs to be adequately evaluated. Does each execution save eight lives, or two, or indeed, any lives?

GUNS AND VIOLENCE

It is worth stating explicitly that there is no logically necessary relationship between levels of gun ownership and interpersonal violence, or even between levels of gun ownership and gun violence. Clearly, most people who own guns do not use them for inflicting harm on others, and many people commit acts of violence without using guns. It is possible for gun violence to decrease while at the same time gun ownership is increasing, as indeed happened in the United States during the 1950s.[2] Gun ownership

[2] Gun ownership increased from an estimated 38,127 guns per 100,000 resident population in 1950 to an estimated 43,061 in 1960, while homicide deaths from firearms declined from 2.75 per 100,000 resident population in 1950 to 2.57 in 1960, and total

can increase among people who are very unlikely to be violent, with or without weapons, while remaining constant among those social groups where interpersonal violence is relatively common, producing no change in the homicide rate. Guns, like other objects, do not have any meaning in and of themselves but only the meaning which is attached to them by people, a meaning which varies between social groups and between social situations. Some people, in some situations, may identify a gun as an instrument of interpersonal violence, a way of establishing dominance, or a method of settling a dispute; other people, in other situations, may not attach any such meaning (Bordua and Lizotte, forthcoming). The mere possession of a gun does not necessarily make deadly violence more likely, and thus the connection between levels of gun ownership and violence is one which must be evaluated empirically.

PREVIOUS RESEARCH

Studies of the deterrent effect of the death penalty and deterrence in general have of necessity been indirect. One cannot observe deterrence since one cannot directly observe nonbehavior, and one can never with any certainty infer that fear of legal sanctions was the reason for someone refraining from a criminal act in any particular situation. Consequently, the range of practical research strategies for studying this problem is quite narrow. The deterrent effect of the death penalty in particular can only be detected indirectly, and in the aggregate, since neither laboratory experimental studies nor self-report survey studies of this subject are practical.

Indirect, aggregate level studies of deterrence (reviewed in Palmer 1977; Tittle and Logan 1973) have all suffered from one or more of several shortcomings.

1. They fail to distinguish between the deterrent and the incapacitative effects of legal sanctions. While the criminal justice system can prevent crime by deterring people through fear of apprehension and punishment, it can also prevent crime simply through the physical restraint of convicted criminals in prisons. Some researchers (Avi-Itzhak and Shinnar 1973; Shinnar and Shinnar 1975) have estimated quite large incapacitative effects, while others (Clarke 1974; Greenburg 1975; Van Dine, Dinitz, and Conrad 1977) find them to be quite modest.

2. Deterrence studies are usually cross-sectional, comparing regions, states, counties, cities, etc., with each other in order to determine the association between the probability and severity of punishment and the

homicides per 100,000 resident population declined from 5.23 to 4.70 during the same period. The computation of these gun-ownership figures is described in the Operationalization section below.

crime rate of these areas (e.g., Ehrlich 1977; Passell 1975). Implicit in these studies is the assumption that the cases are independent, that is, that one area's legal sanctions affect the behavior of residents in that area and, conversely, that one area's residents are influenced only by the legal sanctions of that area and not of neighboring areas. This assumption is especially suspect in any modern society characterized by national communications media and high rates of geographical mobility between legal jurisdictions. For example, of persons in the National Criminal Information Center files arrested for murder between 1970 and 1975, 47% had previously been arrested in other states (U.S. FBI 1975), indicating that they had previously been subject to (and possibly affected by) other states' legal sanctions, not just those of the state in which they committed murder.

3. Some death penalty studies, such as those of Sellin (1967*a*; 1967*b*), examined the relationship between homicide rates and the mere existence of death penalty statutes in American states rather than the actual use of the death penalty. By identifying as "capital punishment states" jurisdictions which had death penalty statutes but had not actually executed anyone for years before the period studied, Sellin biased his investigations toward a finding of no deterrent effect.

4. The majority of previous deterrence studies and almost all of the death penalty deterrence studies (e.g., Schuessler 1952; Sellin 1967*a*) have focused solely on the effect of legal sanction variables on the level of crime, ignoring the effect of the level of crime on the legal sanctions. While increases in the certainty or severity of punishment may reduce the level of crime, increases in the level of crime may reduce the probability or severity of punishment, through an overloading of criminal justice system resources.

5. Probably the most serious shortcoming of deterrence studies, especially death penalty deterrence, is the failure to develop models which would specify variables other than punishment variables which could influence crime rates.

Sellin (1967*a*) attempted to deal with this problem by comparing groups of states which he said were similar in social characteristics but which differed in their having legal provision for the death penalty. However, he did not provide any evidence that the matched states were in fact identical or even very similar in regard to all the characteristics associated with homicide rates, nor could he realistically have been expected to do so, given the large number of such characteristics and the very diverse nature of American states.[3] Other studies making either simple cross-sectional or longitu-

[3] The attempt of Baldus and Cole (1975) to establish the methodological adequacy of Sellin's matching technique backfires on them, since the data which they present on the characteristics of states examined by Sellin, contrary to their interpretation of it, show substantial differences between supposedly matched states, differences which could indeed account for the higher homicide rates of death penalty states (see their table III, p. 178).

dinal comparisons (e.g., Schuessler 1952; Mattick 1966, pp. 10–24; Bailey 1974) suffer from similar problems.

Isaac Ehrlich's recent investigation (1975*b*) provided evidence which contradicted the findings of all the previously cited studies of capital punishment and supported the hypothesis that executions can have the effect of preventing homicides. He avoided many of the errors of previous scholars but had a number of problems of his own. He recognized that it was desirable to extend his time series data as far back in time as possible in order to maximize the number of cases and thus increase the stability of his findings. Unfortunately, he extended the time series so far back (to 1933) that he was unable to obtain the data necessary for estimating a relatively complete model. Data from the FBI *Uniform Crime Reports* (*UCR*), especially on arrest clearance rates and conviction rates, are very unstable and unreliable for most of the 1930s because of the rapidly changing character and number of police jurisdictions supplying reports to the FBI (Bowers and Pierce 1975). Further, extending the time series so far back precluded the inclusion of such variables as availability of firearms, judicial expenditures (for predicting conviction rates), alcohol consumption, and hospital beds (as a measure of the availability of medical care for treating assault wounds), variables for which annual data earlier than the mid-1940s are not available.

Ehrlich also used an inferior measure of his dependent variable, revised *UCR* figures on murders and nonnegligent manslaughter based on police reports instead of vital statistics on deaths due to homicide. The *UCR* data are especially suspect because they were based on incomplete, widely varying collections of jurisdictions, predominantly urban, in the 1930s and early 1940s. In contrast, vital statistics data derive from death registrations, which have been virtually complete since at least 1940 (U.S. Public Health Service 1954). Further, the vital statistics definition of death due to homicide is quite satisfactory for deterrence research purposes since it includes only deaths resulting from injury *purposely* inflicted by other persons (U.S. Public Health Service 1967). Finally, the vital statistics homicide measure is reasonable to use in a time-series analysis since the definition of death due to homicide in the various revisions of the International Classification of Diseases has remained practically unchanged since 1933 (U.S. National Center Health Statistics 1966, p. 9).

Ehrlich's study, like those before his, fails to distinguish between the preventive effects of punishment due to deterrence and the preventive effects due to incapacitation because of imprisonment.[4]

[4] The death penalty may, of course, exert an incapacitative effect on the homicide rate of its own, since those who are executed cannot commit further homicides. However, considering the extremely low rate of homicide recidivism among killers released from

Possibly the most serious flaw in Ehrlich's model was his failure to take account of the possible effect of trends in gun ownership on the homicide rate. Since much of the increase in homicide and other violent crime during this period has been widely attributed to the increasing availability of firearms, it seems essential to attempt to incorporate some aggregate measure of gun ownership in any model of homicide trends, yet in all of Ehrlich's article not a word is devoted to the possible connection between guns and violence.

Although there has been much popular discussion, surprisingly little serious empirical research has studied directly the impact of levels of gun ownership on rates of violent crime. Wolfgang's study of Philadelphia police homicide squad records failed to yield evidence of a causal connection between gun ownership or accessibility and homicide, and the author concluded that few gun homicides "could be avoided merely if a firearm were not immediately present" and that other weapons could be substituted in assaults (1958, p. 83). However Newton and Zimring (1970) claimed to have found that more guns mean more gun violence. Their analysis was methodologically crude, relying on simple zero-order correlations between gun ownership and gun use in violence, thus ignoring other factors which could influence gun use.

Murray's (1975) cross-sectional analysis indicated no relationship between gun ownership and state homicide rates, but Phillips and Votey's (1976) time-series analysis indicated a positive relationship. However, neither study developed any explicit model of homicide causation. Both treated gun ownership as if it were an exogenous variable, even though it is reasonable to expect a simultaneous reciprocal relationship between gun ownership and crime, in which case gun ownership would have to be specified as endogenous. Further, Murray used survey measures of gun ownership which he himself acknowledged to be of doubtful accuracy.

Neither Seitz (1972) nor Fisher (1976) had any direct measure of gun ownership, and both in fact included homicide or assault rates as component items in their gun ownership factors, producing an artifactual association between the rate of violence and their "gun ownership" scores. Further, both authors ignored the possibility of a reciprocal relationship between homicide and gun ownership. Therefore, neither study's findings of a positive relationship between violence and gun ownership can be given much credence.

---

prison (Bedau 1967*b*), it seems very unlikely that executed killers would have killed again even if they had lived. This observation, combined with the fact that few killers are actually executed, suggests that the incapacitative effect of executions on the homicide rate is infinitesimal, whereas the general deterrent effect can theoretically influence almost anyone.

THEORETICAL ORIENTATION AND BASIC MODEL

The model should specify the variables, in addition to gun ownership and deterrence variables, which affect the homicide rate, so that the effects of deterrence and gun ownership can be isolated. Any aggregate model of homicide must focus on "typical" homicides, if there can be such things, rather than just on killings occurring in connection with robberies and other economically motivated crimes, as Ehrlich's model seemed to do. Further, since there are no national data on first-degree murder (the only type of criminal homicide punished by death), but only data on homicide, or murder and nonnegligent manslaughter, it is pointless to develop a model of first-degree murder if it can only be estimated with data referring to all homicides.[5]

What are the major characteristics of homicides? Criminal homicides are most often the result of a domestic quarrel or other argument; a smaller but substantial proportion are committed in connection with other crimes, especially armed robbery (Wolfgang 1958, p. 191; Lundsgaarde 1977, p. 237).[6] Homicides are generally unplanned and impulsive, the result of sudden anger; probably fewer than 5% are premeditated, intentional killings (Wolfgang and Ferracuti 1967, p. 141). A homicide is most likely to occur when alcohol is present in the aggressor (Wolfgang 1958, p. 136). In 1975, 65% of murders and nonnegligent manslaughters were committed with a firearm, 51% with a handgun. Of persons arrested in 1975 for murder and nonnegligent manslaughter, 84.4% were males, 43.7% were persons aged 15 to 24, and 54.4% were blacks. Of the killings, 49% occurred in cities with populations over 250,000 (U.S. FBI 1976).

The subculture of violence thesis of Wolfgang and Ferracuti (1967) provides a useful sociological perspective on homicide.[7] It is sufficiently flexible and comprehensive to serve as the basis for a general theory of social variation in violence and is also consistent with empirically ob-

[5] In Philadelphia from 1948 to 1952 only 15 7% of criminal homicides were designated "first-degree murders" by the courts (Wolfgang 1958, p. 303). In any case, as will be seen later, there is no reason to limit the hypothesized effects of deterrence variables to premeditated homicides only, since the preventive effects of legal sanctions do not necessarily depend on any deliberate, conscious consideration of potential punishments by potential offenders.

[6] In 1975, 32.4% of murders and nonnegligent manslaughters reported to the police were "known or suspected felony killings"; about 54% of *these* were in connection with armed robberies (U.S. FBI 1976, p 19) This was the highest proportion of homicides attributed to felony killings since the FBI started publishing breakdowns of homicides by circumstances.

[7] Space limitations do not permit a review of the subculture of violence perspective or its explanation of the prevalence of violence in these groups. The material is accessible to the general reader, who is referred to Wolfgang and Ferracuti (1967), esp. pp. 95–112, 140–63, and 314–15, regarding the subculture of violence, and pp. 258–64, regarding particular age, sex, class, and race groups.

served patterns of interpersonal aggression in American society. While theoretically appealing, the thesis is difficult to use in empirical research, such as the current research. In an aggregate time-series analysis of U.S. society it is impossible to directly measure strength of commitment to proviolent groups and norms, as there are no annual measures of such a variable. However, an alternative way of operationalizing a subculture of violence perspective in an aggregate level study is to measure membership in those demographic and economic groupings where subcultures of violence are most likely to flourish. Specifically, groupings along four dimensions have been identified as being related to violent behavior and to violent subcultures: age, sex, race, and class (Wolfgang and Ferracuti 1967, pp. 258–64).

From 1947 to 1973 the sex ratio among late adolescents and young adults has been virtually constant, so the young males group is measured simply as persons aged 15–24 per 100,000 resident population.[8] The race grouping is measured as nonwhites per 100,000 resident population, and the lower-class group is measured in income terms as the percentage of families with annual incomes less than $3,000 in constant 1974 dollars.

A subculture of violence is self-perpetuating in the sense that it creates objective conditions of its own which in turn provide justifications for its continuance. It contributes to the dangerousness of any environment where it is prevalent, both for those who are members of the subculture and for those who merely live in the same area. The motivation for acquiring a weapon such as a knife or handgun for self-defense is greater in such an area, making weapon ownership more common, which in turn contributes to the overall dangerousness of the environment. This cyclical process is especially likely to occur where law enforcement is ineffective and where people therefore feel they must rely on their own capacities for self-defense.

In the context of a subculture of violence perspective, the activities of the criminal justice system could be characterized as formal social control efforts to enforce predominantly *anti*violent norms among all social groups, including those culturally committed to relatively *pro*violent norms. The criminal justice system functions to control the level of violence through the mechanisms of deterrence and incapacitation, using formal legal sanctions to minimize the effects of informal proviolent influence.

It has been contended that because of the impulsive nature of homicide one can rule out any deterrent effect of the death penalty a priori (Bedau 1967*a*, p. 271). This line of reasoning relies on the assumption that deterrence requires a conscious, deliberate weighing of risks on the part of the potential offender. This is a very restrictive conception, which has been contradicted by van den Haag, who defined deterrence as a "*preconscious*,

[8] The sex ratio for ages 15–24 was 0.985 in 1940 and 0.981 in 1970 (U.S. Bureau of the Census 1975, p. 15).

general response to a severe but not necessarily specifically and explicitly apprehended or calculated threat" (1969, p. 146, emphasis added). Thus deterrence may depend on a fear of punishment which is neither consciously arrived at nor rationally evaluated.

Given this conceptualization, the deterrent effects of the death penalty need not be limited to premeditated murders but may apply to all homicides, not just those legally defined as punishable by death. The cognitive link in potential offenders' minds may be between the ultimate legal sanction, death, and the act of homicide rather than any particular arbitrary legal subtype of homicide. If this is the case, the appropriate dependent variable in a death penalty deterrence study is the rate of all homicides rather than just those legally defined as capital murders. Therefore the unavailability of statistics on first-degree murder should be no handicap. In any case, given the arbitrary and variable nature of legal definitions, the use of such statistics, if available, would be of questionable value. When Bailey (1974) gathered such statistics for a cross-sectional study, he obtained substantively identical results whether he used first-degree murder or homicide as the dependent variable.

Figure 1 illustrates the preliminary model of U.S. homicide trends. It is a four-equation model, with one equation each for the homicide rate, H, the homicide arrest rate, C, the homicide conviction rate, V, and the estimated per capita rate of firearms, G. Table 1 explains the symbols and indicates the unit of measurement for each variable. The signs on each arrow in figure 1 indicate the predicted direction of the effect of one variable on another. The model as a whole is a dynamic one, including lagged endogenous variables. The inclusion of such variables allows the model to take account of the possibility that endogenous variables (H, G, V, and C in our model) take one or more time periods to adjust to changes in exogenous variables. For example, if the homicide rate adjusts substantially to changes in the percentage of families with incomes less than $3,000 (or any other exogenous variable) over a period of two or three years rather than adjusting immediately, that is, within a single year, then the coefficient of the lagged endogenous variable in the homicide equation, H_{t-1}, should be statistically significant (see Land and Felson [1976, pp. 581–85] for a recent sociological application and discussion of dynamic models). In the initial specification of the model the lagged endogenous variable was included in each of the four equations. The arrest clearance rate, C, and the conviction rate, V, are included in the model as deterrence variables, C being an approximate measure of the probability of the offender being arrested and V being an approximate measure of the probability of the arrested offender being convicted of the homicide offense he was charged with. A simultaneous reciprocal relationship is hypothesized to

exist between each of these variables and the homicide rate: deterrence theory predicts that a higher probability of punishment tends to lower the crime rate, and one could also hypothesize that a high volume of crime tends to put a strain on criminal justice system resources (assuming that publicly allocated resources are not instantly adjusted in accordance with changes in the crime rate) and thus reduces the probability of punishment. Given that none of these three variables (H, C, and V) can legitimately be specified as causally prior to the others, it was appropriate to specify a nonrecursive model.

In addition to these deterrence variables, a measure of incapacitation was included in the model. Although only a very small proportion of the

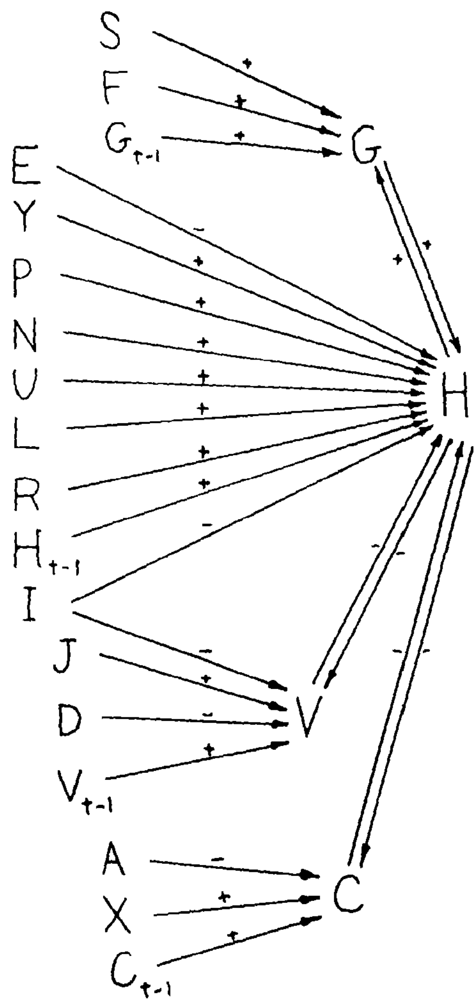


FIG. 1.—Preliminary model of U.S. homicide trends, 1947–73

TABLE 1

VARIABLES USED IN ANALYSIS

Symbol	Variable	Mean (ln)	SD (ln)	Mean (Natural Values)
H	(Homicides – [executions and police homicides])/100,000 resident population	1 706	251	5 696
*G*1	Estimated total firearms/100,000 resident population	10 706	.154	45,198
*G*2	Estimated handguns/100,000 resident population	9 337	213	11,628
C	*UCR* % of murders and nonnegligent manslaughters known to the police cleared by arrest (for total cities)	4 501	045	90 207
V	% of persons charged with murder or nonnegligent manslaughter who are convicted of the offense charged	3 764	088	43 267
*E*1	Executions, year *t*/100,000 estimated persons convicted of murder or nonnegligent manslaughter, year *t*–1	5 463	3 163	1,431
*E*2	Number of executions, year *t*	2 854	1 903	48.111
Y	Persons aged 15–24/100,000 resident population	9 607	.117	14,972
I	State and federal sentenced prisoners present at the end of the year/100,000 resident population	4 678	076	107.82
R	Robberies/100,000 resident population	4 287	472	82 637
P	% of families with incomes less than $3,000 in constant 1974 dollars	2 340	411	11 237
F	Median family income in constant 1974 dollars	9 142	232	9,595
D	Income of families in the twentieth percentile of the income distribution in constant 1974 dollars	8 465	236	4,879
S	Holders of hunting licenses/100,000 resident population	8 981	064	7,968
A	Total state motor vehicle registrations/100,000 resident population	10 626	.212	42,109
J	(State expenditures for judicial activities/price deflator for government purchases of goods and services)/100,000 resident population	6 379	290	616 47
N	Nonwhites/100,000 resident population	9 347	061	11.490
U	Nonfarm residents/100,000 total population	11 413	046	90,581
L	Absolute alcohol, in gallons, consumed per person aged 15+	779	106	2 192
X	(General expenditures on police protection, all governments/price deflator for government purchases of goods and services)/100,000 resident population	9 296	.267	11,280

NOTE.—Variables refer to the United States, each year 1947–73

U.S. population is imprisoned at any one time, imprisonment can nonetheless have a significant incapacitative effect on the homicide rate, independent of its deterrent effects. The prison population is a selective proportion of the total population, consisting heavily of those persons with records of frequent violent behavior in the past and thus especially likely to commit violent acts, including homicides, in the future. The myth that the typical homicide offender is just an ordinary person who slipped once and who had no previous criminal record was effectively laid to rest by Wolfgang (1958, pp. 175–78), who found that 64.4% of his sample of homicide offenders had a previous arrest record, and of these 66.0% had a record of offenses against the person (and arrests undoubtedly underestimate the extent of previous violent behavior). In New York City in 1970, 64.7% of homicide arrestees had a prior arrest record, 40% had prior arrests for violent crimes, and 2% had prior arrests specifically for homicide (Shinnar and Shinnar 1975, p. 596). For the United States as a whole, FBI data indicate that 67.6% of persons arrested for murder or nonnegligent manslaughter from 1970 to 1975 had previous arrest records (U.S. FBI 1976, p. 43). The typical homicide offender could therefore be characterized as a person with a record of previous crimes against the person, any one of which could have resulted in his imprisonment prior to the commission of the homicide.

A simultaneous reciprocal relationship between gun ownership and the homicide rate is specified in the model. It was hypothesized that as gun ownership increases the frequency of homicide increases, for a number of possible reasons. First, firearms are technologically more effective tools for inflicting serious harm (Zimring 1972; Newton and Zimring 1970), and therefore an assault involving a firearm is likelier to result in a fatality than an assault with a less effective weapon like a knife or fists. Second, it is possible that the presence of firearms in certain situations brings about assaults that would not have otherwise occurred at all, as in the case of a small man enabled by the possession of a gun to attack a larger, stronger one.

Looking at the gun-homicide relationship in the other direction, there is good reason to believe that increases in the homicide rate could cause increases in gun ownership by stimulating anxiety and providing greater motivation to obtain a gun for self-defense. Survey data indicate that about 55%–77% of gun-owning households have a gun owned at least partly for protection.[9] Thus the relationship between gun ownership and

[9] In a 1966 NORC national survey, 37% of the households surveyed kept a gun for protection (though it may also have been used for other purposes as well) (Erskine 1972, p. 459). According to a 1965 Gallup national survey, 48% of households had guns (Erskine 1972, p. 456). Assuming this proportion was the same the following year, 0.37/0.48 or about 77% of gun-owning households had a gun for protection. A 1975 national

the homicide rate may be a reciprocal one, with the homicide rate tending to push gun ownership up and gun ownership simultaneously tending to increase the homicide rate.

In addition to the lagged gun ownership rate, G_{t-1}, two other variables are specified in the model as affecting gun ownership. Hunting-license holders per capita, S, is included to account for the sporting use of firearms, since self-defense is by no means the only or even necessarily the primary reason for obtaining and owning guns (Bordua and Lizotte, forthcoming). Median family income (in constant dollars), F, is included to account for trends in the financial capacity to obtain firearms, as distinct from degree of motivation.

Two variables are specified as influencing arrest clearance rates for homicide in addition to the lagged clearance rate, C_{t-1}. Expenditures on police protection, X, are specified to account for the effect which trends in the overall level of police funding may have on the arrest clearance rate. Further, since a huge proportion of police resources is expended on traffic duties, increases in auto ownership, A, leading to increased demand for traffic control, may cause the draining away of police resources from crime-control activity and homicide investigation (Bordua and Haurek 1971).

Judicial expenditures, J, are specified as affecting the conviction ratio, V, since higher funding levels may allow courts to handle more homicide trials and therefore reduce somewhat the necessity for plea bargaining in handling heavy case loads. This could result in more convictions for homicide offenders and in more convictions for the offense charged rather than a lesser offense. Income (in constant dollars) of families in the twentieth percentile of the income distribution, D, is an indicator of the economic resources of the poor, who contribute the bulk of criminal court defendants. It should be inversely related to the conviction ratio since it roughly indexes the capacity of homicide defendants to afford private counsel, instead of having to depend on overworked public defenders or going without a lawyer. The proportion of the U.S. population in prison, I, is specified as having an inverse effect on the conviction ratio since I can be interpreted as a rough measure of prison crowding and because public prosecutors may be more willing to accept pleas of guilty to lesser offenses, which would entail shorter prison sentences at times when prisons are crowded.

The rate of robbery, R, is included in the model to account for homicides which occur in connection with robbery attempts. The proportion of the population not living on farms, U, is used as an imperfect measure (but the only one for which annual data are available) of the proportion of the

survey reported in the *Congressional Record* found that 55% of gun owners had guns at least partially for self-defense (U.S Congress 1976, p. 9).

population which is urban. Finally, *L*, per capita consumption of alcohol, is included in the initial specification of the model to account for the possible connection between aggregate levels of alcohol consumption and the homicide rate.

OPERATIONALIZATION

Table 1 indicates how each variable was measured. The measurement of most of these variables is quite straightforward, but further discussion of some of them is necessary.[10]

The measurement of gun ownership in the United States consisted of cumulating the annual totals of firearms manufactured by domestic manufacturers for private sale in the United States and those imported for private sale in order to estimate the total stock of guns available, and then dividing this figure by the resident population for each year in order to obtain a per capita figure. This procedure was carried out both for total firearms and for handguns alone, as it is conceivable that handgun ownership could be especially important in the explanation of homicide trends since handguns are involved in such a high proportion of homicides. The data have several shortcomings but nevertheless are clearly the best available. Ideally one would like to have reliable data on the percentage of persons or households with a firearm rather than simply the per capita rate, but this kind of data is not available on an annual basis for any substantial number of years. Through 1973 there were national surveys of gun ownership by persons or households in only six years (for these, see Erskine 1972), and their reliability for estimating gun ownership trends is doubtful.[11] The measurement method does not take account of losses of firearms due to destruction, misplacement, or deterioration, but the numbers of such guns over the short period of time studied (1947–73) would in any case be very small compared with the total gun stock. Perhaps the most serious shortcoming of the data used for this measure was that for 1947–68 there was no count of imports of firearm "actions" (trigger and bolt or

[10] Space limitations prevent the inclusion of a complete list of data sources and methods of estimation of missing values. However, the material may be obtained by writing to the author.

[11] For example, perhaps the best national survey of gun ownership through 1973 was the 1968 Harris survey commissioned by the National Commission on the Causes and Prevention of Violence. While the commission's Task Force on Firearms estimated, based on manufacturers' records and importation data, a stock of about 100 million guns in the United States in 1968, estimates derived from the Harris survey indicate only about 66 million guns owned, suggesting that a substantial number of gun owners were lying or "forgetting" about guns they owned. Other national surveys of course faced this same problem.

lever) and other parts which could be assembled into workable guns.[12] It might also be pointed out that data used by the Firearms Task Force to estimate the total of guns manufactured and imported before 1918 were rather incomplete, and therefore the cumulated total subsequent to that period is likely to be an underestimate. However this is not important for our purposes since we are only interested in trends in the stock of guns since 1947.

Two measures of execution risk were used in this study. The first, *E*1, is the number of executions carried out in year *t* divided by the estimated number of persons convicted of murder or nonnegligent manslaughter in the previous year. The figure in the denominator is computed by the following formula: estimated persons convicted of murder or nonnegligent manslaughter $= C/100 \times V/100 \times$ total homicides committed, where *C* and *V* are as defined in table 1. Thus, *E*1 is a true measure of execution risk since it approximates the proportion of persons convicted of a homicide offense in a year who were executed in the following year (the mean lag between conviction and execution during the period studied was about a year [Lunden 1962, p. 1043; U.S. Bureau of Prisons 1970]). The second, *E*2, is simply the number of executions in year *t*. If one assumes that executions exert a deterrent effect by "reminding" people of the existence of the penalty and that the perceived risk of suffering the death penalty is therefore greater in years when executions are more frequent, then *E*2 is a more appropriate measure than *E*1. Ideally, one would of course like to have a direct measure of the perceived risk of execution for each year, but of course there is no available aggregate measure of the perceived risk, either of execution or of arrest or conviction for homicide. Thus one must settle for using the objective risks as surrogates for the perceived risks.

In theoretically underdeveloped areas such as deterrence study or criminology in general, theory is rarely sophisticated enough to specify the exact mathematical form of the relationships between variables. Most statistically estimated models in sociology are simply stated in ordinary linear additive form, or occasionally in log-linear form. There are at least two a priori reasons to prefer a log-linear specification, where the variables are expressed as natural values, although neither of them constitutes decisive or indisputable justification. First, as Ehrlich (1975*a*, pp. 217–18) has suggested, if one assumes that the magnitude of the errors in measurement of a variable is proportional to the level of that variable, the log-linear tends to minimize the influence of measurement error. Second, the natural log transformation of variables emphasizes proportional changes or differences in the level of variables. This seems to be a desirable quality for theoreti-

[12] For other limitations of the data, see Newton and Zimring (1970, pp. 171–72). Data on gun production and importation after 1968 were gathered by the U.S. Bureau of Alcohol, Tobacco, and Firearms, and covered "actions."

cal reasons if one believes, for example, that a change in the execution risk from 100 executions per 100,000 homicide convictions to 200 per 100,000 is more significant in impact on the homicide rate than a change from 1,800 executions per 100,000 convictions to 1,900 per 100,000. Ultimately, the decision between functional forms has to be settled empirically, given the limited theoretical development in this problem area. The model was estimated using both specifications of functional form in order to determine which specification results in a better fit of the data.[13]

The model was estimated on annual U.S. data for the period 1947–73.[14] It was estimated using two-stage least squares rather than ordinary least squares because a nonrecursive model, with simultaneous reciprocal relationships between endogenous variables, was specified. Application of ordinary least squares to such a model could result in biased and inconsistent estimates of the regression coefficients (see Johnston [1972, pp. 341–420] for an explanation of simultaneous-equations methods and Mason and Halter [1968] for an early example of the use of two-stage least-squares estimation in sociology).

RESULTS

The preliminary model illustrated in figure 1 was estimated and a number of the predicted relationships were found to be insignificant; therefore several of the variables initially included in the model were deleted (S, hunting license holders per capita; V_{t-1}, the conviction ratio lagged one year; X, expenditures on police protection; E, execution risk; N, nonwhites per 100,000 resident population; U, nonfarm residents per 100,000 resident population; L, alcohol consumption, and H_{t-1}, the homicide rate lagged one year were dropped). Figure 2 illustrates the resulting basic model and table 2 contains the estimated parameters of the model's structural equations. In the last two columns of table 2 four descriptive statistics are given: D-W, the Durbin-Watson statistic, is a measure of autocorrelation of the disturbances in an equation (for a discussion of auto-

[13] Theil (1971, p. 543–45) recommends use of the variance of the residual term of an equation estimated under both competing specifications as a criterion for choosing between functional forms where there is no solid a priori theoretical justification for one particular specification.

[14] While it is generally desirable to maximize one's sample size, the time series for this study could not be extended back further than 1947. It was pointed out in the literature review that the source of much of the Ehrlich study's shortcomings was the extension of the time series beyond the point where adequate reliable data were available. For the model presented here, annual data on variables P, F, D, J, and L are not available before 1947 and data on I, M, U, S, J, and D disappear around 1944–46. Perhaps most crucial of all, annual data on the production and importation of firearms are not available before 1945.

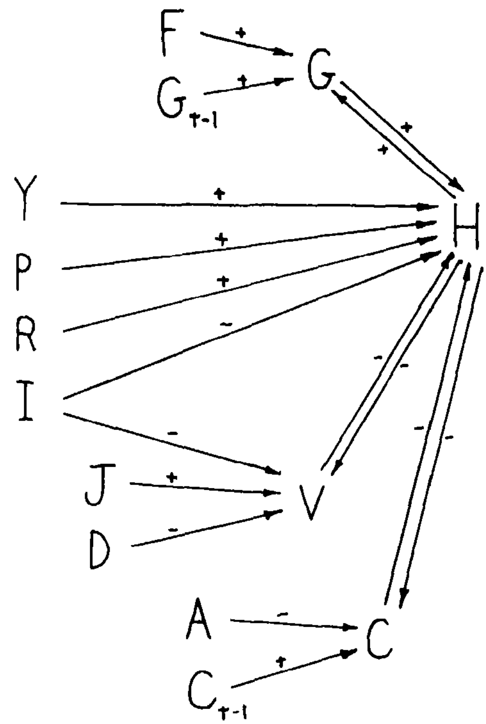


FIG. 2.—Final model of U S. homicide trends, 1947–73

correlation and the D-W statistic, see Johnston [1972, pp. 243–66]); SEE is the standard error of estimate, that is, the standard deviation of the error term of the equation; df is the degrees of freedom of the equation; and Σe_t^2 is the sum of the squared residuals, a measure of the fit of the equation to the data.[15]

Equations 1–4 in table 2 constitute the model of U.S. homicide trends, gun ownership trends, and related criminal justice system activity, using *G*1, total firearms per 100,000 resident population, as the measure of gun ownership. Equations 5–8 constitute the same model, except that *G*2, handguns per 100,000 resident population, is the gun variable.[16] The

[15] The reader will note that the adjusted R^2 statistic conventionally reported in connection with ordinary least-squares estimation is not reported here. This is because the R^2 statistic of structural equations estimated in two-stage least squares is not directly relevant or meaningful for such equations and cannot be interpreted in the straightforward way that the R^2 in ordinary least squares is. In fact, the structural equation R^2 can even take on negative values (Bassman 1962). The adjusted R^2 (corrected for degrees of freedom) in the first stage of the homicide equations reported in this paper was invariably at least .97.

[16] Total guns is the sum of longguns (rifles and shotguns) and handguns. Both longguns and handguns per capita cannot be included in the same equation because of their high collinearity. (Their zero-order correlation is 0.97.)

TABLE 2

UNSTANDARDIZED STRUCTURAL EQUATION COEFFICIENTS FOR COMPLETE MODEL
(Ratio of Coefficients to Their SEs in Parentheses)

Equation No	Dependent Variable	Intercept	$\ln \hat{G}1$	$\ln \hat{C}$	$\ln \hat{V}$	$\ln Y$	$\ln I$	$\ln R$	$\ln P$	D-W df	SEE Σe_i^2
1	$\ln H$	729	404	−1 547	− 247	685	− 686	149	229	2 35	01900
			(2 62)	(−5 98)	(−2 11)	(6 97)	(−5 73)	(2 75)	(6 28)	19	00686
			$\ln \hat{H}$	$\ln G1_{t-1}$	$\ln F$						
2	$\ln G1$	476	061	921	032					1.16	00505
			(8 81)	(33 13)	(2 15)					23	00059
			$\ln \hat{H}$	$\ln C_{t-1}$	$\ln A$						
3	$\ln C$	1 983	− 053	675	− 040					1.78	01269
			(−1 84)	(3 54)	(−2 64)					23	00370
			$\ln \hat{H}$	$\ln J$	$\ln I$	$\ln D$					
4 .	$\ln V$	9 166	− 715	521	− 806	− 441				1 89	05613
			(−2 85)	(1 90)	(−1 73)	(−2.01)				22	06930

Equation No	Dependent Variable	Intercept	$\ln \hat{G}2$	$\ln \hat{C}$	$\ln \hat{V}$	$\ln Y$	$\ln I$	$\ln R$	$\ln P$	D-W df	SEE Σe_i^2
5	$\ln H$	932	367	−1 326	− 153	622	− 686	145	212	2 34	01835
			(2 27)	(−4 14)	(−1 37)	(7 27)	(−5 60)	(2 41)	(6 54)	19	00640
			$\ln \hat{H}$	$\ln G2_{t-1}$	$\ln R$	$\ln D$					
6 .	$\ln G2$	308	026	861	045	092				1 62	00683
			(1 11)	(25 69)	(2 37)	(5 49)				22	00102
			$\ln \hat{H}$	$\ln C_{t-1}$	$\ln A$						
7	$\ln C$	2 053	− 056	659	− 040					1 75	01268
			(−1 94)	(3 47)	(−2 62)					23	00370
			$\ln \hat{H}$	$\ln J$	$\ln I$	$\ln D$					
8 .	$\ln V$	9 263	− 726	532	− 824	− 449				1 89	.05612
			(−2 92)	(1 95)	(−1 78)	(−2 05)				22	.06928

model contains four equations, one for each of the endogenous variables: H, the homicide rate; G, gun ownership; C, the homicide arrest clearance rate; and V, the homicide conviction rate. As an aid to interpreting the tables, equation 1 is written out in full:

$$\ln H = 0.729 + \underset{(2.62)}{0.404 \ln \hat{G}1} - \underset{(-5.98)}{1.547 \ln \hat{C}} - \underset{(-2.11)}{0.247 \ln \hat{V}} + \underset{(6.97)}{0.685 \ln Y}$$

$$- \underset{(-5.73)}{0.686 \ln I} + \underset{(2.75)}{0.149 \ln R} + \underset{(6.28)}{0.229 \ln P} + e_H$$

where ln denotes the natural logarithm of each variable, e_H is the residual for the homicide equation, the coefficients are unstandardized coefficients, and the variables are as defined in table 1. The number in parentheses below each regression coefficient is the ratio of the coefficient to its standard error.[17]

The coefficients presented in table 2 are consistent in direction with the predictions of our theoretical model. Gun ownership, whether measured as total guns or handguns, has a significant positive effect on the homicide rate. Both the arrest clearance rate and the conviction rate show significant negative effects, supporting a general deterrence hypothesis. Further, the measure of incapacitation, I, shows a significant negative effect on the homicide rate independent of the clearance and conviction rates, supporting the hypothesis of an incapacitative effect of imprisonment on the homicide rate. The reader will note that none of the execution risk measures appear in table 2; this is because they generally were found to be not significantly related to the homicide rate. I will discuss this later.

The age structure variable, Y, the percentage of poor, P, and the robbery rate, R, all had significant positive effects on the homicide rate, as predicted. Coefficients estimating the effect of the homicide rate on either total gun ownership or handgun ownership are in the predicted direction, although the homicide coefficient in the handgun-ownership equation (6) is a bit low, possibly because handgun ownership also responded to the robbery rate, another cause of the acquisition of guns for defensive reasons. Total guns per capita lagged one year, G_{t-1}, and median family income, F, had positive coefficients in the total guns equation. Judicial expenditures, J, showed a positive effect on the conviction rate, while income level of the poor, D, showed a negative effect, again as predicted. The homicide clear-

[17] With ordinary least-squares estimation, this ratio would be a t-ratio, but since there are no small sample properties defined for the two-stage least-squares estimator, the ratio is assumed to have a normal distribution and is therefore called a z-ratio. A z-ratio of 1 645 or higher is significant at the 5% level (one-tailed), while a value of 1.282 is significant at the 10% level.

ance rate lagged one year, C_{t-1}, had a positive coefficient, and autos per capita, A, had a negative coefficient in the clearance-rate equation, again as predicted.

No significant relationship was found between per capita police expenditures and the homicide clearance rate. Also, neither the number of policemen per capita nor the number of *UCR* "index" offenses (other than homicides) per policeman, a measure of the competition for police resources from other felonies, showed any significant effect on the arrest clearance rate.

The variables indicated in figure 1 which do not appear in figure 2 were deleted because they had insignificant regression coefficients. The sizes and significance levels of estimated coefficients were interpreted conservatively, and so no variable was deleted from the model simply because it was not significant at the arbitrary .05 level. The choice whether or not to delete a variable was rarely difficult, however, as the deleted variables almost always had coefficients with a probability of occurring by chance greater than .20. Most of these null results are not reported. However, a few of the more interesting and/or important ones are reported in table 3.

Table 3 can be interpreted in the same way as table 2 except that only the homicide equation's parameters under various specifications are reported, as it is these parameters which are of principal substantive interest. In any case, the parameters of the other equations remain qualitatively unchanged in estimations of various specifications of the model. Equation 1 in table 3 is the same as the homicide equation in the basic model described in table 2, except that proportion of the population nonwhite, N, was included. Its coefficient, reported in the Additional Variables column, was insignificant at even a one-tailed .10 level, possibly because N, like many other nonsignificant variables, had very little variance over the period studied. In equation 2, proportion of the population which was nonfarm, U, a surrogate for proportion urban, had an insignificant coefficient, as did per capita alcohol consumption in equation 3. In addition, a number of other variables were included (one at a time) in the specification of the homicide equation and found to have insignificant coefficients, including a medical care index,[18] the unemployment rate, the percentage of the population in the civilian labor force, and both the unemployment rate and

[18] A medical care index was created, using iterated principal axis factor analysis and varimax rotation, based on four items: hospital beds/100,000 resident population, hospital personnel/100 patients (for general hospitals), hospital admissions/100,000 resident population, and total M.D.'s/100,000 resident population. The index could be considered primarily a measure of the availability and use of medical care. It therefore did not tap improvements in medical technology and their diffusion, improvements such as new techniques in surgical repair, fluid replacements, prevention of blood loss, and control of shock. Such improvements are relevant to the treatment of assault wounds and could therefore influence the homicide rate (Wolfgang and Ferracuti 1967, pp. 294–95).

TABLE 3

UNSTANDARDIZED STRUCTURAL EQUATION COEFFICIENTS FOR HOMICIDE EQUATIONS
(Ratio of Coefficients to Their SEs in Parentheses)

Equation No	Intercept	ln $\hat{G}1$	ln $\hat{C}$	ln $\hat{F}$	ln Y	ln I	ln R	ln P	Additional Variables	D-W df	SEE Σe_i^2
1	−7 380	083	−1 861	− 305	622	− 906	147	384	1 546	2 33	02023
		(26)	(−4 76)	(−2 12)	(5 18)	(−3 81)	(2 54)	(2 69)	(1 13)	18	.00736
2	2 591	427	−1 344	− 165	689	− 624	173	192	322	2 30	01910
		(1 79)	(−1 34)	(− 61)	(5 56)	(−1 44)	(1 69)	(69)	(− 137)	18	00657
3	−1 969	567	−1 461	−2 723	824	− 747	.148	210	− 450	2 49	.01878
		(2 86)	(−5 67)	(2 30)	(5 81)	(−5 83)	(2 78)	(5 43)	(−1 29)	18	00635
4	975	596	−2 225	− 331	894	− 875	183	362	− 300	2 07	02090
		(2 54)	(−3 51)	(−2 09)	(4 62)	(−4 27)	(2 99)	(3 18)	(−1 26)	18	00786
5	532	405	−1 509	− 234	691	− 706	153	218	003	2 31	01931
		(2 57)	(−4 32)	(−1 79)	(6 53)	(−4 28)	(2 69)	(3 32)	(43)	18	00671
6	937	432	−1 427	− 272	649	− 859	199	184	016	2 43	01795
		(2 94)	(−5 69)	(−2 46)	(6 79)	(−5 62)	(3 66)	(4 35)	(1 78)	18	00580
7	1 317	408	−1.790	− 316	712	− 540	081	295	− 027	2 37	01889
		(2 66)	(−5 78)	(−2 44)	(7 29)	(−3 75)	(1 18)	(5 79)	(−1 79)	18	00642
8	518	424	−1 524	− 266	687	− 698	147	230	001	2 38	01972
		(2 21)	(−5 66)	(−1 89)	(6 72)	(−4 97)	(2 53)	(5 95)	(46)	18	00700
9	1 178		−1 332	192	661	− 510	337	170		1 98	03300
			(−2 37)	(59)	(4 00)	(−2 69)	(3 78)	(4 15)		20	02178
10	540	310	−1 635		694	− 618	209	224		1 98	02309
		(1 64)	(−4 86)		(5 82)	(−4 29)	(3 68)	(5 04)		20	01066
11	− 142	678	−1 830	− 336	778	− 848		251		2 21	02404
		(4 92)	(−5 30)	(−2 50)	(6 70)	(−6 63)		(5 65)		20	01156
12	14 607	369	−2 772	− 542	037	− 872	132	360		2 16	02951
		(1 54)	(−6 90)	(−2 98)	(24)	(−4 69)	(1 57)	(6 38)		19	01654
13*	12 450	00008	− 122	− 041	0002	− 025	0072	124		2 17	13118
		(2 97)	(−5 01)	(−1 96)	(4 57)	(−3 23)	(1 61)	(6 01)		19	32697
14	−2 869	607	−1 190	− 219	686	− 745	133	248		2 30	.01860
		(1 94)	(−2 21)	(−2 13)	(6 95)	(−5 28)	(2 39)	(5 35)		17	00588
15	491	290	−1 223	− 234	671	− 686	186	199		2 42	.01971
		(1 25)	(−2 74)	(−1 99)	(6 19)	(−5.44)	(2 60)	(3 50)		17	.00660
16	1 067	453	−1 553	− 289	664	− 765	123	230		2 28	01594
		(3 72)	(−7 54)	(−3 12)	(8 59)	(−7 95)	(2 94)	(7 79)		19	00573

NOTE —See text for specific additional variables referred to in each equation.
* Variables in natural values

the labor force participation rate together. The latter two variables showed no significant effect on the homicide rate, although Ehrlich (1975*b*) found such an effect. This may have been because of the inclusion, in the present study, of the robbery rate in the homicide equation, suggesting that unemployment and labor force participation only affect the homicide rate insofar as they influence the commission of economically motivated crimes, which in turn may produce homicides as an unintended consequence. In equation 4, the homicide rate lagged one year, H_{t-1}, had an insignificant coefficient in the homicide equation (in the Additional Variables column), indicating that the homicide rate was adjusting fairly completely to exogenous forces within a one-year period.

Equations 5–8 illustrate the principal results regarding the deterrent effect of the death penalty. Each equation is the basic homicide equation, with a different measure of execution risk included. Equations 5–8 include $E2_t$, $E1_t$, $E2_{t-1}$, and $E1_{t-1}$, respectively. The estimated coefficients do not provide any consistent support for the hypotheses that the death penalty exerts any unique deterrent effect on homicides. The coefficient for $E2_t$ is insignificant and positive, while that for $E1_t$ is significant and positive. Since it could be argued that people do not respond to current sanction risks, but to risks prevailing in the immediate past, the model was also estimated with $E2_{t-1}$ and $E1_{t-1}$ in the homicide equation. The coefficient of $E2_{t-1}$, in equation 7, was insignificant and negative, consistent with a deterrence hypothesis, but in equation 8, $E1_{t-1}$ had an insignificant positive coefficient. Furthermore, when execution variables with lags of two to five years were used, they all had insignificant coefficients, both for $E1$ and $E2$. (These findings are consistent with the insignificant coefficient for H_{t-1}, indicating that the homicide rate adjusts to exogenous forces within a single year.) Finally, it should be mentioned that if execution risk is treated as an endogenous variable in the model, as if it had a simultaneous reciprocal relationship to the homicide rate, its coefficient in the homicide equation is still not significantly different from zero.[19]

Given the small number of cases and the multicollinearity typical in aggregate time-series analyses, it is prudent to test the stability of the estimated parameters of the model with regard to the sample period studied, the functional form of the model (e.g., whether linear or nonlinear relationships are specified), the variables included in the model, and the measures of the variables used. We have already discussed a number of vari-

[19] A *simultaneous* reciprocal relationship, with the homicide rate in year t influencing the number of executions in year t, is in any case unlikely since executions in year t are to a great extent foreordained in year $t-1$, when the death sentences were handed down. Such a relationship could exist only if the homicide rate in year t significantly and consistently influenced the number of pardons, commutations of sentence, and stays of execution in year t.

ables which might have been included in the model but were found to be irrelevant, and we have discussed the various measures of execution risk used. Equations 9–11 are different versions of the homicide equation, each with one of the following variables omitted from the model: *G*1, *V*, and *R*. The results of omitting these three variables are presented because they are probably the three variables in the model most suspect with regard to measurement error. The omission of *G*1 in equation 9 leaves all of the coefficients qualitatively unchanged except that the coefficient of *V* becomes positive and insignificant. This was not felt to be sufficient cause for omitting *V* from the model, as its coefficient is in accord with theoretical expectations in other specifications of the model. In equation 10, *V* was omitted, leaving the coefficients of the other variables substantially unchanged. In equation 11, *R* was omitted, again leaving the other variables' coefficients qualitatively unchanged. The omission of other variables in the model produces similar results.

Although it was pointed out earlier why the *UCR* measure of the homicide rate is inferior to that derived from vital statistics, the model was nonetheless estimated using the *UCR* measure of the dependent variable in order to see if it would produce drastically different results. The results in equation 12 are quite similar to those obtained using the vital statistics measure. The one exception is that the coefficient of *Y* becomes insignificant. The reader should also note that the sum of squared residuals (Σe_i^2) of the homicide equation with the *UCR* measure is more than twice as large as that with the vital statistics measure (equation 1 in table 2), indicating an inferior fit of the model with the *UCR* measure, possibly because of substantial measurement error in the *UCR* figures.

No consistent relationship was found between executions and the homicide rate, regardless of functional form. Further, when the basic model, without execution risk, was estimated with the natural values of the variables (linear relationships rather than log linear), qualitatively identical results were obtained. Consistent with Theil's (1971) recommendation of the residual variance criterion for choosing between functional specifications, I compared the sum of squared residuals for the two specifications. A direct comparison between equation 13 in table 3 and equation 1 in table 2 is not possible since the dependent variable is measured in different units in the two equations, so the residuals resulting from estimation with the log-linear specification were converted into natural values to make them comparable with the residuals from the natural value (linear) specification.[20] When the conversion was performed, the sum of squared

[20] The procedure is described in more detail in Theil (1971, p. 544). If the residuals from the natural value specification are converted into natural logs (the converse of the procedure just described), the log linear specification is still found to be distinctly superior to the natural value specification.

residuals from the log-linear specification was 0.215, compared with 0.327 for the natural value specification (in equation 13). Thus it seems that the natural value functional form produces a distinctly inferior fit with the data.

In a time-series analysis it is of some value to see how dependent one's results are on a specific set of time points. While it is generally desirable to maximize the number of time points in the sample, by deleting a number of observations one can produce various subsamples of years. Then the model can be estimated on pairs of subsamples with the same number of time points, to see if the parameter estimates differ substantially between the two subsamples. Since there are only 27 time points to work with between 1947 and 1973, it is pointless to delete more than two of them, since the deletion of even as many as three years leaves only 14 degrees of freedom in the reduced-form equations, leading to unstable parameter estimates regardless of the adequacy of the model or the data. In equations 14 and 15 two years have been deleted from the sample. Equation 14 was estimated for the period 1949–73, while equation 15 was estimated for 1947–71. The two equations are comparable with each other, both having the same number of degrees of freedom, and produce essentially similar results. When the model was estimated on the periods 1947–72 and 1948–73 (i.e., one year deleted), virtually identical results were produced.

When a simultaneous-equations system shows nonzero contemporaneous covariances between structural equation disturbances, and at least one of the equations is overidentified, three-stage least-squares (3SLS) estimation can produce more efficient parameter estimates (Theil 1971, p. 511). As these conditions obtained in the present situation, 3SLS estimation was applied to our final model to see if the 2SLS results were confirmed. Equation 16 is the homicide equation estimated with 3SLS. For every coefficient, the z-ratio was higher for the 3SLS estimate, and standard errors of the coefficients were generally substantially lower. Therefore, the more efficient 3SLS estimates confirmed the results obtained with 2SLS.

In addition to estimating a proposed model under various alternative specifications, another way of evaluating its adequacy is to see how well it forecasts values of the endogenous variables for time points in advance of those on which the model was estimated, conditional on observed values of the predetermined variables.[21] Since my model was developed using data for the period 1947–73, forecast values were obtained for H, G, C, and V, by substituting the 1974 values of the predetermined variables into the reduced-form equations for each of the endogenous variables. The resulting forecast values and forecast errors are shown in table 4. The forecasts

[21] Land and Felson (1976) provide a sociological example of forecasting with a dynamic structural equation model and discuss its use in somewhat greater detail than provided here.

TABLE 4

FORECASTS OF 1974 VALUES OF ENDOGENOUS VARIABLES BASED ON REDUCED-FORM EQUATIONS ESTIMATED FOR 1947-73

Variable	ln *H*	ln *Gi*	ln *C*	ln *V*
Forecast value	2.2701	11.0962	4.3850	3.5968
Observed value	2.3003	11.0937	4.3808	3.7087
Error	−.0302	.0025	.0042	−.1119
% error	1.31	.02	.10	3.02
SEE of equation	.01737	.00301	.00770	.04568
Error/SEE	−1.74	.83	.55	−2.45

are generally quite accurate, with less than 1½% error, except for *V*, the conviction rate. This exception is not surprising considering the relatively small population on which reports of court conviction rates are based, the changing composition of this subpopulation, and the generally erratic *UCR* reporting procedures for this measure. An exact test of the accuracy of these forecasts would involve computation of the standard errors of the forecasts, but as these are always greater than or equal to the standard errors of the equations, one can say that all the forecasts are within three standard errors of forecast, and all but one are within two standard errors. The forecasts are thus consistent with the hypothesis that the observed 1974 values of the endogenous variables were generated by the same causal structure as that estimated for 1947–73.

DISCUSSIONS AND CONCLUSIONS

We have seen that the basic model is fairly robust with regard to the variables included or excluded, the measures of the dependent variable and the key independent variables used, the functional form of the model, the sample of the time points included, and the estimation procedure used. Further, the model is consistent with individual-level sociological theory and research on violent behavior and formal and informal social control. Nevertheless, several limitations of the study and its findings should be mentioned. The study necessarily had to be done at an aggregate level, making it impossible to get measures of the subjectively perceived risk of legal punishment. I essentially had to assume a roughly constant proportionality over time between perceived and objective risks of punishment. If one assumes that the use of objective risks as imperfect surrogate measures of perceived risks produces random measurement error, then the effect of using such surrogates is to bias the deterrence variables' coefficients to zero. This tends to strengthen confidence in the findings regarding clearance and the conviction rates but to weaken confidence in the finding of no preventive effect of the death penalty. Another limitation is the

omission of any measure of the severity of noncapital punishments for homicide, essentially the average prison sentence served or handed down. It was not possible to obtain or estimate any reliable time series on sentence length since only fragmentary data exist for this measure, either at the state or the national level.

Because of data limitations, it was not possible to reliably estimate any adequate model on a sample of time points extending further back than 1947. As the 1930s and early 1940s were a time of high homicide rates and also relatively high execution rates, it seems that the inclusion of these years in the sample would have contributed to producing a positive correlation between execution risk and the homicide rate, which would work against the death penalty deterrence hypothesis and thus strengthen our findings of no deterrent effect. However, it could quite plausibly be suggested that some minimal, threshold level of execution risk must be achieved before any deterrent effect of the death penalty will be felt, and further that this threshold level was not generally reached during the period 1947–73, a period of quite low execution risks. Therefore it is not possible to state definitively that the death penalty never has exerted any deterrent effect in the United States, or that it never could. It also is not clear that the death penalty cannot deter specific relatively rare but generally premeditated crimes such as kidnapping or skyjacking. Nevertheless, the evidence we have examined does not support the hypothesis that the death penalty exerted any general deterrent effect on the homicide rate in the United States during the period since World War II On the other hand, there was no consistent support for the idea that executions increase homicide, either through a "brutalizing" effect, diminishing respect for human life, or by stimulating imitative violence, as has been suggested by previous writers (e.g., Sellin 1959, pp. 65–69).

The more general deterrence hypothesis relating the probabilities of arrest and conviction for homicide and the frequency of acts of homicide was consistently supported. Contrary to the recent finding of Logan (1975), even homicide, impulsive though it usually is, seems to be deterrable by legal sanctions, even though it is not specifically deterred by the rarely imposed death penalty.

Imprisonment was shown to have an incapacitative effect, distinct from deterrence effects. The simple physical constraint of imprisonment, imposed on a particular segment of the population, appears to exert a significant preventative effect on the homicide rate.

The findings of this study regarding the death penalty and its alleged deterrent effect contradict those of Ehrlich's recent study (1975*b*). One possible reason for these contradictory findings would be a misspecification of his model by Ehrlich, specifically his omission of one of the more obvi-

ously important explanatory variables: gun ownership. That portion of the increase in the U.S. homicide rate in the 1960s which Ehrlich's model would attribute to the decline in executions may well have been caused instead by rapid increase in the stock of privately owned firearms. Evidence was found for the hypothesis of a deadly circularity of gun ownership and homicide, homicide pushing up gun ownership and gun ownership at the same time pushing up homicide. These findings must be qualified somewhat by observing that it is quite possible that it was increases in gun ownership in traditionally violent subgroups of the population, rather than simply the overall increase, which contributed to the homicide increases during the 1960s and 1970s.

The aggregate level association found in this research between the homicide rate and the level of gun ownership, combined with individual level survey data on reasons for obtaining and owning firearms, strongly suggests that crime is a cause of gun ownership just as gun ownership is a cause of crime. If people own guns because they fear crime, because they see themselves in serious danger of being assaulted, raped, burglarized, or robbed at gun point, then the task of reducing violence through the reduction of gun ownership may be a profoundly difficult one. As long as crime remains at a high level and people feel they have to depend on their own resources for self-defense they will continue to acquire guns for protection and to retain those they already own. Short of politically impractical or undemocratic strategies such as forced confiscation of guns, effective programs aimed at reducing gun ownership may be doomed unless a reduction in crime can be made first, or until people at least have sufficient confidence in the ability of the criminal justice system to protect them from victimization so that they would not feel they had to depend on themselves and on their guns for protection.

REFERENCES

Avi-Itzhak, Benjamin, and Reuel Shinnar. 1973. "Quantitative Models in Crime Control." *Journal of Criminal Justice* 1 (Fall): 185–217.

Bailey, William C. 1974. "Murder and the Death Penalty." *Journal of Criminal Law, Criminology, and Police Science* 65 (September): 416–21.

Baldus, David C, and James W. L. Cole. 1975. "A Comparison of the Work of Thorsten Sellin and Isaac Ehrlich on the Deterrent Effect of Capital Punishment." *Yale Law Journal* 85 (December): 170–86.

Bassman, R. L. 1962. "Letter to the Editor." *Econometrica* 30 (October): 824–26.

Bedau, Hugo Adam. 1967*a*. *The Death Penalty in America*. New York: Doubleday.

———. 1967*b*. "Parole of Capital Offenders, Recidivism, and Life Imprisonment." Pp. 395–405 in Bedau 1967*a*.

Bordua, David J., and Edward W. Haurek. 1971. "The Police Budget's Lot." Pp. 57–69 in *Police in Urban Society*, edited by Harlan Hahn. Beverly Hills, Calif.: Sage.

Bordua, David J., and Alan J. Lizotte. Forthcoming. "Patterns of Legal Firearms Ownership: A Cultural and Situational Analysis of Illinois Counties."

Bowers, William J., and Glenn Pierce. 1975. "The Illusion of Deterrence in Isaac Ehrlich's Research on Capital Punishment." *Yale Law Journal* 85 (December): 187–208

Clarke, Stevens H 1974 "Getting 'em out of Circulation. Does Incarceration of Juvenile Offenders Reduce Crime?" *Journal of Criminal Law, Criminology, and Police Science* 65 (December): 528–35.

Ehrlich, Isaac. 1975*a*. "Deterrence: Evidence and Inference." *Yale Law Journal* 85 (December): 209–26.

——— 1975*b*. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *American Economic Review* 65 (June)· 397–417

———. 1977. "Capital Punishment and Deterrence Some Further Thoughts and Additional Evidence." *Journal of Political Economy* 85 (August): 741–88

Erskine, Hazel. 1972. "The Polls: Gun Control." *Public Opinion Quarterly* 36 (Summer): 455–69.

Fisher, Joseph C. 1976. "Homicide in Detroit· The Role of Firearms." *Criminology* 14 (November): 387–400.

Greenburg, David F 1975. "The Incapacitative Effect of Imprisonment " *Law and Society Review* 9 (Summer): 541–80

Johnston, John 1972. *Econometric Methods* New York: McGraw-Hill

Land, Kenneth C , and Marcus Felson. 1976 "A General Framework for Building Dynamic Macro Social Indicator Models Including an Analysis of Changes in Crime Rates and Police Expenditures " *American Journal of Sociology* 82 (November)· 565–604.

Logan, Charles H 1975. "Arrest Rates and Deterrence " *Social Science Quarterly* 56 (December). 376–89.

Lunden, Walter A. 1962. "Time Lapse between Sentence and Execution The U S and Canada Compared " *ABA Journal* 48 (November)· 1043–45

Lundsgaarde, Henry P 1977. *Murder in Space City: A Cultural Analysis of Houston Homicide Patterns* New York: Oxford University Press

Mason, Robert, and Albert N. Halter 1968 "The Application of a System of Simultaneous Equations to an Innovation Diffusion Model " *Social Forces* 47 (December) 182–95

Mattick, Hans W 1966 *The Unexamined Death* Chicago John Howard Association

Morris, Norval, and Gordon Hawkins 1977. *Letter to the President on Crime Control.* Chicago: University of Chicago Press

Murray, Douglas R. 1975 "Handguns, Gun Control Laws, and Firearms Violence " *Social Problems* 23 (October)· 81–93

Newton, George D , and Franklin Zimring 1970 *Firearms and Violence in American Life* Washington, D C.· Government Printing Office

Palmer, Jan. 1977 "Economic Analyses of the Deterrent Effect of Punishment· A Review." *Journal of Research in Crime and Delinquency* 14 (January)· 4–21

Passell, Peter. 1975 "The Deterrent Effect of the Death Penalty A Statistical Test " *Stanford Law Review* 28 (November) 61–80

Phillips, Llad, and Harold L. Votey, Jr. 1976 "Handguns and Homicide: Minimizing Losses and the Costs of Control." *Journal of Legal Studies* 5 (June)· 463–78

Schuessler, Karl F. 1952 "The Deterrent Effect of the Death Penalty." *Annals* 284 (November): 54–62.

Seitz, Steven T. 1972 "Firearms, Homicides, and Gun Control Effectiveness." *Law and Society Review* 6 (May): 595–614.

Sellin, Thorsten 1959 *The Death Penalty*. Philadelphia· American Law Institute

———. 1967*a*. "Death and Imprisonment as Deterrents to Murder " Pp 274–84 in Bedau 1967*a*.

———. 1967*b*. "Effect of Repeal and Reintroduction of the Death Penalty on Homicide Rates." Pp. 339–43 in Bedau 1967*a*.

Shinnar, Reuel, and Shlomo Shinnar. 1975 "The Effects of the Criminal Justice System on the Control of Crime: A Quantitative Approach." *Law and Society Review* 9 (Summer): 581–611.

Theil, Henri. 1971. *Principles of Econometrics*. New York: Wiley.

Tittle, Charles R., and Charles H. Logan. 1973. "Sanctions and Deviance: Evidence and Remaining Questions." *Law and Society Review* 7 (Spring): 371–92.

U.S. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970*. Washington, D.C.: Government Printing Office.

U.S. Bureau of Prisons. 1970. *National Prisoner Statistics*. Bulletin no. 45. Washington, D.C.: Government Printing Office.

U.S. Congress. 1976. *Congressional Record*. Vol. 121, no. 129, pt. 2.

U.S. Federal Bureau of Investigation (FBI). 1933–76. *Uniform Crime Reports*. Washington, D.C.: Government Printing Office.

U.S. National Center Health Statistics. 1966. *Homicide in the U.S., 1950–64*. Washington, D.C.: Government Printing Office.

U.S. Public Health Service. 1954. *Vital Statistics of the United States 1950*. Washington, D.C.: Government Printing Office.

———. 1967. *International Classification of Diseases*. 8th rev. Washington, D.C : Government Printing Office.

van den Haag, Ernest 1969. "On Deterrence and the Death Penalty " *Journal of Criminal Law, Criminology, and Police Science* 60 (June)· 141–47.

Van Dine, Stephan, Simon Dinitz, and John Conrad. 1977. "The Incapacitation of the Dangerous Offender· A Statistical Experiment." *Journal of Research in Crime and Delinquency* 14 (January): 22–34.

Wolfgang, Marvin E. 1958. *Patterns in Criminal Homicide* Philadelphia: University of Pennsylvania.

Wolfgang, Marvin E., and Franco Ferracuti. 1967. *The Subculture of Violence: Towards an Integrated Theory in Criminology*. New York: Tavistock.

Zimring, Franklin E. 1972. "The Medium Is the Message: Firearm Caliber as a Determinant of Death from Assault." *Journal of Legal Studies* 1 (January): 97–123.

Education and Political Party: The Effects of College or Social Class?

Terry S. Weiner
Union College

Bruce K. Eckland
University of North Carolina at Chapel Hill

The relationship between education and political party preference is examined for nonsouthern males and females using retrospective data from a national sample of high school students followed up in 1970 at about age 30. As expected, the zero-order correlation between years of schooling and Republican party preference was found to be positive. However, controlling for inherited partisan loyalties, class origins, and socioeconomic attainment, the positive correlation disappeared and, at least for males, the direct effect of higher education on being Republican was found to be negative, as those who view college as a socializing agent would have predicted. The implications of these findings for the apparent decline and future of class-based politics, as the proportion of college-educated voters continues to increase over the next several decades, are briefly discussed.

A paradox regarding the relationship between education and political party preference has existed for many years in the literature on political socialization and voting behavior. Put briefly, the theory that colleges and universities function as important socializing agents in the development of tolerance of ambiguity and a belief in civil liberties and equality would not normally lead one to hypothesize a positive correlation between educational attainment and identification with the "conservative" Republican party. However, in most national and regional studies on voting behavior, except in the South, just such a relationship is typically observed (Knoke 1972; Pomper 1975).

This paradox, we believe, can be resolved only if political socialization is viewed as a continuous process over the life cycle and if college is viewed not simply as a socializing agent but also as a credentialing agent in the general socioeconomic attainment process. Our central thesis is that the persistent positive correlation between educational attainment and being a Republican does not reflect the direct effects of higher education but can be accounted for by two social processes unrelated to whatever direct impact college has on political values.

The first of these is the selective recruitment into higher education of persons from middle- and upper-middle-class Republican backgrounds, coupled with the tendency for party preferences to be acquired early in life from parents and carried over into adulthood because of inertia. Thus, if both educational attainment and adult partisan loyalties have common antecedents rooted in the family of origin, the association between going to college and voting Republican could be entirely spurious.

A second set of conditions which could produce a positive correlation between years of schooling and identification with the Republican party is that after the completion of formal schooling people tend to adjust partisan loyalties to match those of the dominant status groups to which they belong and in which membership depends on educational attainment. As a credentialing agent, higher education is a powerful determinant of occupational status and social mobility through marriage. If socialization continues after college and conservative political-economic values are associated with high-status jobs and upper-middle-class status or, conversely, if the Democratic party is the worker's party, educational attainment again would be expected to correlate with being a Republican.

Thus, it should not be surprising that, apart from any direct effects that educational contexts may have on changing values, the combined effects of inherited partisan loyalties, class background, and the credentialing or social mobility function of education are likely to be at least as important and to contribute to the observed correlation.

An accurate assessment, then, of the effects of education on party preference requires testing a causal model in which both destination and background measures of social class are included and, preferably, parents' party preference as well. This will be the primary objective of the analysis presented in this report. The data are based on a 1955 national Educational Testing Service (ETS) sample of high school sophomores who were followed up in 1970 at about age 30 by the Institute for Research in Social Science at the University of North Carolina. The subsamples used in this analysis include 504 males and 474 females, excluding students from schools located in the South.

THE MODEL

The zero-order correlations between educational attainment and a Republican party preference for the respondents in our survey are small, although in the same direction as found in other studies, .03 for males and .12 for females. The corresponding correlations for their fathers' education and party, as reported by the respondents, were .27 and .33, respectively. An

explanation for the difference between the two generations and its implications will be offered at the end of the report.

Our goal here is to test three alternative hypotheses concerning the relationship between education and party, two of which, as already mentioned, predict a positive correlation between years of schooling and identification with the Republican party. The third hypothesis, that college changes values, predicts a negative correlation. Since we are anticipating offsetting or suppressor effects, the size of the observed correlations is not of immediate importance. In fact, if the zero-order coefficient were zero, the issue being addressed here still would be relevant. However, since the signs of the education-party correlations for both males and females are positive, we will be particularly interested to see if the signs can be reversed when appropriate prior and intervening test factors are controlled.

Hypothesis One: Early Political Socialization and Selective Recruitment to College

There is some debate over whether the family actually transmits the kinds of attitudes and learnings that would lead to logically coherent political ideologies or whether partisanship among children is largely devoid of specific information about what the parties stand for (Hess and Torney 1967; Hyman 1959; Jennings and Niemi 1968; Dawson and Prewitt 1969; Jaros 1973). Nevertheless, there is broad agreement that party choice, like the selection of a particular religious denomination, is determined by the family and is well formed by the time a child reaches adolescence. In fact, so strong are inherited partisan loyalties that it is virtually a truism in political sociology that the best single predictor of party preference among adults, aside from spouse's preference or how one voted in the past, is how one's parents voted when one was growing up (Knoke 1972).

Moreover, politics historically have been class based. There are, of course, many known ethnic, religious, urban-rural, and regional differences (Knoke 1976; Hamilton 1972; Converse 1966; Matthews and Prothro 1964), and, too, it appears that the relationship between social class and voting behavior has been declining over the past generation or more (a topic we will come back to later). Nevertheless, the working class still tends to vote Democratic, while the middle and upper-middle classes tend to vote Republican (Abramson 1974; Alford 1963; Campbell 1960; Lipset 1960). For purposes of the immediate discussion, how this pattern happened to evolve is unimportant. What is significant here is the correlation itself and the fact that educational attainment is strongly dependent on class background (Blau and Duncan 1967; Spady 1967; Alexander, Eckland, and Griffin

1975). As a consequence, students who graduate from high school and enter college are recruited disproportionately from Republican backgrounds. Given, in addition, the relative stability of inherited party preference with the passage of time, we would expect this set of correlations to produce a positive but spurious association between educational attainment and identification with the Republican party in adulthood.

The first step in conceptualizing our model then (see fig. 1) is to introduce socioeconomic background and the political party preference of one's parents as exogenous and correlated variables. Educational attainment is shown to depend on class origins, while the respondent's current party preference depends, among other things, on his parents' party preference. Since class origins and parents' party (Republican) are positively correlated, educational attainment and the respondent's preference will be correlated to some extent and in the same direction.

Hypothesis Two: Social Mobility and Adult Resocialization

Apart from whatever effects educational institutions may have on changing values and beliefs, they act as important credentialing agents in the allocation of social status, which, in turn, places people into positions which may lead to further changes in values and beliefs. More specifically, the two steps in this process are (*a*) individuals who earn college and advanced degrees are placed in occupations or other status groups that traditionally have tended to be Republican and (*b*) these same individuals subsequently are resocialized to take on the life-styles and political orientations of the status groups in which they now hold membership and the political attitudes characteristic of the roles they now occupy. Political socialization,

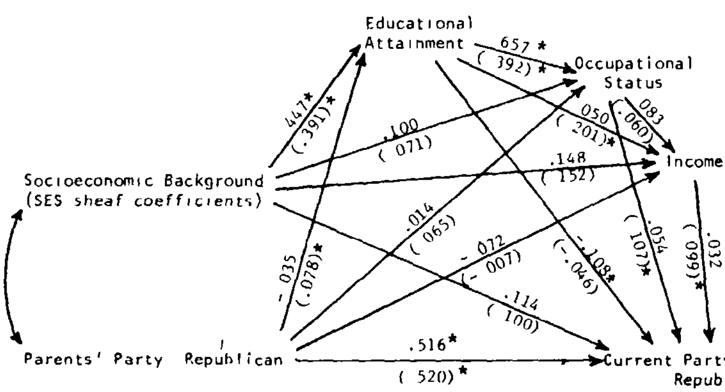


FIG. 1.—Basic model of the effects of education in the political socialization process. (All figures are standardized regression coefficients, those in parentheses are for females, and those marked by asterisks are more than twice their standard error.)

then, does not end with leaving home or school but is a lifelong process. Changes in a particular party preference are likely to be associated with an individual's indoctrination into new social roles and changes in his position in the class structure. These two steps are discussed further below.

The correlation between educational and occupational attainment is remarkably strong (Alexander, Eckland, and Griffin 1975; Blau and Duncan 1967), basically, we believe, because of both the extensiveness of economic and professional specialization (Trow 1961; Eisenstadt 1956) and the importance that various status groups assign to educational certification as a mechanism for limiting and maintaining control over their membership (Collins 1971). In other words, educational institutions provide students with both the skills and the credentials typically required for entry into elite positions in the occupational hierarchy and other status groups. In the case of women, this includes mobility through marriage (Elder 1969).

Where young people end up as members of different status groups, in turn, is likely to affect their political outlook. As previously noted, different occupational and social classes tend to share common party preferences. The higher a person's status or position is, generally speaking, the more likely he is to support the conservative (Republican) party, that is, the one that promises to maintain the status quo on economic issues. On the other hand, labor unions and working-class voters tend to support the liberal party, that is, the one that promises to reduce economic inequalities (Alford 1963; Rose 1974). The correlation between social class and political party preference in the United States, however, is relatively weak and possibly becoming weaker. This is not only because of the cross-pressures of race, ethnicity, and region but also because on most noneconomic issues, like civil liberties, abortion, and school desegregation, the working classes are less liberal than the middle and upper-middle classes (Stouffer 1955; Hamilton 1972). Nevertheless, on economic issues, such as medical care and guaranteed incomes, the correlations between politics and class generally are consistent and almost always are found to be in the predicted direction. Low social status is associated with being more liberal and being a Democrat, while high social status is associated with being more conservative and being a Republican.

What this suggests, then, is that college students from lower-status Democratic families for whom education was a vehicle for upward mobility are likely to shift their party loyalties after college to match those shared by others in their class of destination, whether for reasons of cognitive consistency, peer group pressure, or economic rationality (McClosky and Dahlgren 1959; Lipset 1959; Dawson and Prewitt 1969). Although the findings of past studies are not entirely consistent, there is evidence that mobile individuals who move from a social environment where most people

consider themselves Democrats to one which is predominantly Republican tend in time to undergo a change in party preference from the one appropriate to their original status to that typical of their new environment and status (Barber 1970; Maccoby, Matthews, and Morton 1954; Thompson 1971). Similarly, we expect that students from middle- and upper-middle-class Republican backgrounds who may have shifted their party loyalties to a more liberal position during college are likely to shift back again to the conservative orientations of their parents as they once again move in social circles that are predominantly Republican.

Although the researchers did not take the question of mobility into account, a series of follow-ups conducted by the National Opinion Research Center of alumni who graduated from college in 1961 found an upswing for the Republicans of 5% and a downswing for the Democrats of 6% between 1964 and 1968 (Greeley and Spaeth 1970). Thus we predict that, independent of the effects of one's class origins, the status allocation function of higher education would lead indirectly to a positive correlation between educational attainment and identification with the Republican party. To test this hypothesis, we entered occupational status and income into our model as intervening variables, temporally located between educational attainment and current party preference.

Hypothesis Three: Changing Values in College

Institutions of higher education, of course, not only place people in the class structure but also presumably change people. It is quite possible, however, that the contextual effects of college on political values and partisan loyalties may be completely masked by the offsetting effects of family background and other status variables on the education-party correlation. The argument to be developed here is that one important direct consequence of higher education is, in fact, its inducement of "liberal" changes in party preference.

Most colleges and universities provide considerably more freedom for self-examination and criticism of society than other social institutions, including the family and work place as well as earlier schooling. In elementary and secondary education, instilling political attitudes involves deliberately training children in citizenship and the development of consensual values and an attachment to the nation and government (Hess and Torney 1967; Coleman 1965; Dawson and Prewitt 1969). In contrast, colleges and universities generally reject the role of in loco parentis and instead provide some separation and protection for faculty and students devoted to critical assessment and to shaping new ideas and values that often are at the leading edge of society and divergent from those of previous generations (Clark 1962; Keeton 1971).

There is considerable evidence that college generally does liberalize attitudes. Although some institutions, particularly the private liberal arts colleges, have distinctive ideologies and therefore no doubt differ widely in their impact on students (Knoke and Isaac 1976), and although even within a particular college environment the impact will not be uniform for all students, the net change is in the liberal direction. This appears to be true on most measures of political liberalism, such as tolerance for ambiguity, openness to change, a belief in equality, and humanitarian attitudes (see Feldman and Newcomb [1970] for a review of the literature).

With the exception of college polls (which are unreliable because they are cross-sectional) and Newcomb's study of Bennington college girls in the 1930s (Jaros 1973), there is no direct evidence that students actually change their party preference during the college years. Nevertheless, most institutions of higher education, even those which are church controlled, are at least to some degree committed to free inquiry and innovation and are relatively detached from the political mores of the general culture. Since the mid-20th century, their normative climates have been breeding grounds for liberal philosophies.

It also could be argued that most colleges, particularly public institutions, contain a wider variety of social backgrounds than the neighborhoods and high schools from which students come (McConnell and Heist 1962). Independent of their parents often for the first time, entering freshmen confront a social environment marked by a degree of diversity of interests and beliefs previously unknown to many of them. In such a situation, their predispositions and the sometimes ethnocentric attitudes they have acquired from their parents are likely to be tested, challenged, and, to some extent, changed (Webster, Freedman, and Heist 1962; Clark et al. 1972), often without much deliberation or coaching. The end result, as most authors suggest, is to increase the students' tolerance of diversity and their readiness to live in a society without racial, ethnic, or economic barriers.

There are several factors, then, that would lead us to hypothesize that higher education, at least in the short run, will tend to break down the partisan loyalties of students, especially those from conservative Republican backgrounds, and incline students generally toward more liberal, and sometimes radical, party choices.

METHODS

The data used for this study are part of a nationally representative survey of public high school sophomores conducted by ETS in 1955 and a follow-up survey conducted by the Institute for Research in Social Science at the University of North Carolina in 1970. The follow-up sample included

947 male and 1,130 female respondents. Detailed descriptions of the schools and the sampling procedures are available elsewhere (Eckland and MacGillivray 1972; Alexander and Eckland 1973, 1974). For purposes of this report, all respondents from schools in the South have been excluded because of the ambiguity of the relationship between liberalism-conservatism and party identification in a region that still tends to be dominated by one party. This restriction, plus missing data on some specific items, reduced the sample to 504 men and 474 married women.

The model in figure 1 was analyzed separately for males and females because of two important differences in the measures employed, as explained below. The beta weights shown in the path diagram are net standardized regression coefficients, with those for females presented in parentheses. All variable measures in this analysis were obtained from the 1970 schedule.

Socioeconomic background.—Four status background measures are included in the analysis. The father's occupation is taken from a precoded Edwards-type scale, which was collapsed from nine to six status rankings, and which is his occupation when the respondent was in high school. Both the father's and the mother's education are from a set of nine precoded response categories ranging from "none, or some grade school" to "Ph.D. or professional degree." The fourth measure is a household index of possessions in the respondent's parents' home "when you were growing up." An initial list of 22 items was reduced to a factor-weighted index of 13 household commodities. The index may be considered a surrogate measure of parental income.

Since we are interested not in the separate but in the aggregated effects of these measures, we have employed sheaf coefficients to summarize the influence of an unmeasured socioeconomic construct (SES). Computationally, the sheaf coefficient involves summing the direct status regression coefficients for a given dependent variable on the four background measures and then correcting that sum to take into account the interrelationship among the status indicators themselves (Heise 1972). The sheaf coefficient for SES is interpreted as a standardized coefficient and may be used in the same manner as other regression coefficients to estimate total, direct, and indirect effects.

Party preference.—Party identification was determined from a question which asked about the respondent's own political leanings "at present" and those of his or her parents "15 years ago," that is, when the respondent was in high school. Response categories were combined to form a four-point scale ranging from Democrats (low), through liberal independents and conservative independents, to Republicans (high). The placement of independents here is similar to the procedure used by Knoke (1972), who

found that using a Republican–non-Republican dichotomous measure of party preference did not yield different results. Finally, a small number of respondents who reported "other" party preferences were excluded from the analysis.

We make no claim that this variable, as constructed, is a direct measure of political liberalism or conservatism, although we do assume that Republicans are more conservative, on the average, than Democrats despite the various constituencies each party serves. This can be justified on two bases: the perceptions of the parties by the public and the actual stance of party members on political issues. On the first point, recently Pomper (1975) reported that approximately two-thirds of both working and middle classes still perceive the Democrats as more liberal than the Republicans. On the second point, Nie, Verba, and Petrocik (1976) reported that in 1972 significantly more Republicans characterized themselves as conservative than as liberal, while the reverse was true for Democrats.

It also will be noted that our measure of parents' party relies on both retrospective data and the perceptions of the offspring and not on the parents themselves. Niemi (1973) has shown that measures of parents' party identification obtained from students are only moderately reliable (about .60), in which case the effects of parents' party in our model would tend to be underestimated. On the other hand, a quite different problem when using recall data, as Niemi also demonstrates, is that respondents tend to claim a closer correspondence between their own party preference and what they perceive to be their parents' party than is actually the case, which would tend to inflate rather than deflate these same correlations in our data. Thus, it appears that the effects of these two sources of error would tend to offset each other. However, even if they did not, we have no reason to expect that measurement error is an unusually serious problem here.

Educational attainment.—This measure is identical with the scale used for the father's and the mother's education, but it applies to the respondent and has seven categories, ranging from "some high school" to "Ph.D. or professional degree."

Occupational status.—This measure is identical with the scale used for the father's occupation, but it applies to the male respondents. In the case of females, the husband's occupation was used. (For this reason, never-married females were excluded from the analysis.)

Income.—"Personal income" was derived for males from a 24-point scale ranging from "under $2,000" to "$36,000 and over." For females, "total household income" was used instead because of the relatively low labor force participation rates for women in this age group and the low wages of those who were employed. We presume for this age cohort that

the husband's earnings and, thus, the total consumption power of the family is more important than the wife's earnings for studying the consequences of economic status on party identification.

The zero-order correlations between all variables included in the model are presented separately for males and females in table 1, with respondents from the South excluded from the analysis. Our presentation will focus initially on the findings for the male sample, and then we will discuss those for the females.

Males

With regard to the zero-order correlations, three brief comments are in order. First, the association between the male respondent's party preference (Republican) and educational attainment is small but positive, as noted earlier. Second, parents' party affiliation also is positively related to each of the four indices of class background, including education, and the correlations are all of about the same magnitude. And third, of the various correlates of the respondent's party preference at age 30 included in the model, the party of his parents when he was growing up, not surprisingly, is by far the strongest (.54 for males).

We now move to the regression coefficients and the three hypotheses outlined earlier. The first of these, the selective recruitment hypothesis, stated that part of the observed correlation between educational attainment and political party preference is spurious because of the effects of class background on who goes to college, the association between parents' party and educational attainment, and the stability of the parent-offspring party correlation over time. This hypothesis receives fairly strong support. First, the son's educational attainment is indeed associated with his parents' political party preference in the predicted direction (.11 in table 1), meaning that students from Republican backgrounds are more likely to graduate from high school and go to college. The correlation, as expected, is entirely a function of the effects of class background on educational attainment and the covariance between class and political party (see fig. 1). And, controlling for class background, educational attainment, occupation, and income, the partial correlation between parents' and son's party at age 30 remains strong, the coefficient being reduced only from .540 to .516. Using the tracing rules of path analysis, the structural network which links educational attainment to class origins, to parents' political party, and to the son's party, would alone produce a .099 correlation between educational attainment and a Republican party preference in 1970, as shown in table 2.

Let us next examine the social mobility hypothesis. The theory, as stated

TABLE 1

ZERO-ORDER CORRELATIONS AMONG ALL VARIABLES

	Father's Occupation	Father's Education	Mother's Education	Household	Parents' Party	Education	Occupation	Income	Party	Mean	SD
Father's occupation ..		43	26	30	27	30	28	14	15	2 79	1 77
Father's education	52		50	42	27	36	30	12	22	3 35	1 75
Mother's education	.32	48		35	24	22	22	11	18	3 63	1 35
Household index..	42	48	47		23	32	28	13	19	8 27	2 84
Parents' party .	29	33	28	33		11	12	00	54	2 50	1 46
Education.	32	33	30	34	23		69	16	03	5 50	1 74
Occupation..	18	16	09	15	16	41		16	08	3 64	1 71
Income	17	16	13	.22	11	28	17		03	10 74	4 00
Party.	.10	18	10	.12	51	12	17	14		2 48	1 34
Mean	2 66	3 50	3 57	8 50	2 52	4 87	3 66	11 80	2 51		.
SD..... .	1 67	1 94	1 45	2 93	1 47	1 32	2 24	4 61	1.35	. .	.

NOTE —Coefficients above the diagonal are for males ($N = 504$), those below the diagonal are for females ($N = 474$)

TABLE 2

DECOMPOSITION OF OBSERVED CORRELATIONS BETWEEN EDUCATION AND PARTY PREFERENCE

Basic Components	Males	Females
Selective recruitment (spuriousness due to family background)	099	102
Social mobility (indirect effects mediated through occupation and income)	.039	064
Changing values in college (the direct effect of education). .	−.108	− 046
Totals (observed zero-order correlations between education and Republican party preference). . .	.030	120

earlier, is that higher educational attainment leads to higher-status jobs and higher income which, in turn, place college graduates in a sociopolitical context that leads them to support the more conservative party, thereby reversing the party preferences of students who upon graduation may have been leaning away from the Republican party. As figure 1 shows, the first set of links in this causal chain is reasonably strong for males, at least with regard to job attainment. The direct effect of education on occupational status is .657, while the total effect of education on income is equivalent to a coefficient of .105. (The latter is the sum of the direct effect, .050, and the indirect effect via occupation, .657 times .083.)

However, the second set of links in this process, occupation and income to party preference, does not lend much credence to the resocialization hypothesis. Although the coefficients for males are in the predicted direction, they are too small to be statistically significant. As a consequence, the total indirect effect of educational attainment on political party preference, via both the son's occupational status and income, is only .039 (see table 2).

It appears, then, that the early socialization hypothesis linking education and current party preference back to one's family origins, and therefore spuriously to each other, contributes more to the tendency for college-educated men to identify with the Republican party than does the social mobility hypothesis. The combined effect of the two sets of equations derived from hypotheses one and two nevertheless produces a .138 correlation between education and party preference. The fact that the observed correlation, in contrast, is only .039 is strong evidence of the existence of suppressor effects, which brings us to our final and perhaps most important hypothesis.

According to the third hypothesis, the direct effect of educational attainment on Republican party preference was expected to be negative when the offsetting effects of class background and socioeconomic attainment are controlled. The path coefficient in figure 1, which shows the direct effect

of education to be a significant —.108 for males, clearly supports this hypothesis. It is the only negative path to current party preference in the model. Higher education, then, appears to have substantively important political consequences as a socializing agent. Independent of their class origins and the partisan loyalties they inherited as adolescents, students who go to college and graduate are less likely to identify with the Republican party than those who do not go to college or never finish high school.

Females

The above findings may be generalized to apply to females, with some exceptions. We have examined, but not reported here, the unstandardized regression coefficients or metric values, and in no instance do they alter the conclusions drawn below. The reader also should be reminded before proceeding that the female model differs in two ways. Spouse's occupation is used instead of the respondent's own occupation, whether she was employed or not, and total household income replaces personal income.

To begin with, the observed correlation between educational attainment and Republican party preference, as noted at the outset of this article, is stronger for females (.12) than for males (.03). Does this mean that in comparison with men, higher education has a smaller liberalizing effect on the partisan loyalties of women or simply that the effects of family origins and mobility through marriage are stronger in determining their party preferences? The answer is developed below.

a) As noted in table 2, the total effects of family background on the education-party correlation are virtually identical for men and women, .099 and .102, respectively. Moreover, the specific manners in which the background variables operate in the model do not particularly differ, except in one minor respect. Regardless of class background, women from Republican families were somewhat more likely than others to complete high school and go to college. This correlation was not found for men. Nevertheless, the overall impact of early socialization and selective recruitment on the education-party correlation is the same for women and men.

b) The social mobility or resocialization hypothesis receives somewhat stronger support from the findings for women than for men. The specific process also differs, although this is not surprising, since we have substituted husband's occupation for the respondent's and total household income for personal income in the female model. In the first set of paths, the beta weight between a woman's educational attainment and her husband's occupation is not as strong as that between a man's educational attainment and his own occupation (see fig. 1). On the other hand, the beta weight between educational attainment and household income is stronger

for women than is the direct effect of education on personal income for men. Taken together, it appears then that college is about as important an avenue for social mobility through marriage for women as it is an avenue for mobility for men in terms of their own socioeconomic achievements.

The more relevant sex differences, for our purposes, appear in the second set of paths, the effects of occupational status and income on party preference. Both of the direct paths are slightly stronger for women than for men. The partisan loyalties of women, in other words, appear to be more a direct function of their husband's occupational class and household income than are the loyalties of men a function of their own job and income. Therefore, when we take both sets of paths into account, we find that the total indirect effect of educational attainment on the party preferences of women is slightly larger than that found for men, .064 versus .039.

c) The direct effect of higher education on Republican partisan loyalties also is in the predicted direction for women (negative), but the coefficient is not statistically significant. College, it appears, has less direct impact on changing the party preferences of women than of men.

In conclusion, the stronger positive zero-order correlation between educational attainment and Republican party preference for women than for men is due to a combination of the somewhat stronger indirect effects of the social mobility function of education in remolding the partisan loyalties of women and the somewhat weaker direct effects of college as a socializing agent in breaking down the political predispositions of women.

DISCUSSION

Using retrospective data collected in 1970 on Northern subsamples of males and females taken from a national longitudinal survey of high school sophomores who were followed up at about age 30, the effects of education on political party preference were examined in the context of a status attainment model. As anticipated, the zero-order correlations between educational attainment and Republican party preference were found to be slightly positive for both men and women. These correlations, however, do not accurately reflect the direct effects of formal education on political values or party identification. For both males and females, the socialization effects of higher education are masked to some extent by the status allocation function of higher education but more significantly by the fact that going to college is associated with both class origins and the inherited partisan loyalties of one's parents. When family origins and class of destination are controlled, the partial correlation between educational

attainment and a stated preference for the Republican party becomes negative.

In our opinion, the data clearly answer the question with which we began. That is, despite the observed relationship between higher education and conservative partisan politics typically found in voting studies, education has a net liberal influence on a person's political party preference—in fact, it is the only such factor in our model. Yet, we are left with the question, Is the influence of this factor large enough to have any part in the apparent decline of class-based politics in the United States? We think so.

First, let us briefly reexamine the two conditions probably most important to the persistence of class-based politics: (*a*) the transfer simultaneously of class status and political party loyalties between generations through the family and (*b*) the effects of social mobility on partisan loyalties. Either party loyalties and status are carried down jointly from parents to offspring or the class-party correlation must be developed anew each generation. With regard to the first process, much less than half of the variance in occupation, income, or the respondent's party preference is explained in our model by family background The strongest intergenerational effect of any of the background variables is that of parents' party on their offspring's party. However, even this factor explains only about 30% of the variance in current party preference. Thus, it would not seem that inherited party loyalties could carry for very long the weight of class-based politics from one generation to the next.

The only other possibility would be a strong correlation between status attainment and political party. Yet, our analysis suggests that, controlling for family background, the direct effects of occupational status and income on party preference are really very weak. Moreover, as noted at the beginning of this report, the observed positive correlation between educational attainment and Republican party preference is considerably weaker today than it was for the parents of our respondents. It is not at all clear, however, why this is so. One possible reason is the continuing rise in educational levels combined with the probable net effect of higher education in breaking down old party loyalties.

Between 1940 and 1960 the proportion of all 18–21-year-olds enrolled in college more than doubled (Trow 1961). Even if this trend does not continue and the figures stabilize, the average educational level of the adult population nevertheless will continue to rise for at least the next 4 or 5 decades, just as it has for the past several decades. This opens up the possibility that higher education is strongly implicated in the decline of class-based politics. And there is a related point.

In contrast to the rise earlier in this century of mass public secondary education, which by design still functions largely to conserve the common

culture and the moral authority of the state, higher education functions in almost the opposite way. It is not a "melting pot" but consists of several thousand relatively autonomous and diverse institutions, both public and private. There probably are as many different kinds of private liberal arts colleges each with its own distinctive character as there are religious and ethnic groups in this country. And even the large public universities contain within their separate boundaries numerous student subcultures and status groups of various political bents. That colleges differ in their impact on the sociopolitical attitudes of students is well established (Knoke and Isaac 1976).

Given this kind of diversity in higher education and the ever-increasing proportion of the adult population exposed to its mix of cultural relativism, secular values, particularism, universalistic standards, and humanism, we should not be surprised if political partisan loyalties are more difficult to sustain today than in the past. Neither of the two major parties today offers an internally consistent and logically coherent ideology. How long they will continue to be class based is an open question.

Certainly, party identification and American political behavior will continue to be based on social cleavages. It seems, however, that the substance and form of these cleavages is changing and becoming less class based, in part, we believe, because of the rise of mass higher education. These new cleavages, as Weber might have predicted, are more likely to reflect professional and bureaucratic interests and to be more differentiated and delimited in scope than those found in traditional class conflict politics (Janowitz and Segal 1967). The endurance of the two national parties and their ability to support diverse constituencies that cut across class boundaries will continue to be tested as successive generations of more highly educated and professionally oriented men and women enter the voting-age population.

REFERENCES

Abramson, P. 1974. "Generational Change in American Electoral Behavior." *American Political Science Review* 68 (March): 93–112.

Alexander, K. L., and B. K. Eckland. 1973 "Effects of Education on the Social Mobility of High School Sophomores Fifteen Years Later (1955–1970)." Final Report, U.S. Office of Education, Project no. 10202 (DEE-4-71-0037).

———. 1974. "Sex Differences in the Status Attainment Process." *American Sociological Review* 39 (October): 668–82.

Alexander, K. L., B. K Eckland, and L. J. Griffin 1975 "The Wisconsin Model of Socioeconomic Achievement: A Replication." *American Journal of Sociology* 81 (September): 324–42.

Alford, R. 1963. *Party and Society: The Anglo-American Democracies.* Chicago: Rand McNally.

Barber, J. A. 1970. *Social Mobility and Voting Behavior.* Chicago: Rand McNally.

Blau, P M., and O. D. Duncan. 1967. *The American Occupational Structure* New York Wiley

Campbell, A , P E Converse, W. Miller, and D. Stokes 1960. *The American Voter.* New York· Wiley.

Clark, B. R. 1962 *Educating the Expert Society* San Francisco· Chandler

Clark, B. R., D. Herst, T McConnell, M Trow, and G. Yonge 1972 *Students and Colleges: Interaction and Change.* Berkeley. Center for Research and Development in Higher Education

Coleman, James S , ed 1965 *Education and Political Development* Princeton, N J Princeton University Press

Collins, Randall. 1971 "Functional and Conflict Theories of Educational Stratification." *American Sociological Review* 36 (December) 1002–19

Converse, P E 1966 "Religion and Politics The 1960 Election " Pp 96–124 in *Elections and the Political Order,* edited by A Campbell et al New York Wiley

Dawson, R E , and K. Prewitt 1969. *Political Socialization* Boston· Little, Brown

Eckland, B. K., and L MacGillivray 1972 "School Profiles " Working Paper no 1. Chapel Hill, N C Institute for Research in Social Science.

Eisenstadt, S. N 1956 *From Generation to Generation: Age-Groups and Social Structure* Glencoe, Ill : Free Press

Elder, G 1969 "Appearance and Education in Marriage Mobility " *American Sociological Review* 43 (August) 519–33

Feldman, K A., and T M Newcomb 1970 *The Impact of College on Students* San Francisco Jossey-Bass

Greeley, A M , and J L. Spaeth 1970 "Political Change among College Alumni " *Sociology of Education* 43 (Winter) 106-113

Hamilton, R. F 1972 *Class and Politics in the United States* New York· Wiley.

Heise, D R 1972 "Employing Nominal Variables, Induced Variables, and Block Variables in Path Analysis " *Sociological Methods and Research* 1 (November)· 147–73

Hess, R D , and J V Torney. 1967 *The Development of Political Attitudes in Children.* Chicago: Aldine.

Hyman, H. 1959. *Political Socialization* Glencoe, Ill · Free Press

Janowitz, M , and D R Segal 1967 "Social Cleavage and Party Affiliation Germany, Great Britain, and the United States " *American Journal of Sociology* 72 (May)· 601–18

Jaros, D 1973 *Socialization to Politics* New York. Praeger

Jennings, M K , and R. G Niemi 1968 "The Transmission of Political Values from Parent to Child " *American Political Science Review* 62 (March)· 169–84

Keeton, M T 1971 *Models and Mavericks* New York McGraw-Hill.

Knoke, D 1972 "A Causal Model for the Political Party Preferences of American Men." *American Sociological Review* 37 (December). 679–89

———. 1976. *Change and Continuity in American Politics* Baltimore· Johns Hopkins University Press

Knoke, D , and L Isaac. 1976 "Quality of Higher Education and Sociopolitical Attitudes." *Social Forces* 54 (March)· 524–29

Lipset, Seymour Martin 1959. *Political Man.* Garden City, N.Y Doubleday.

McClosky, H , and H. E Dahlgren 1959 "Primary Group Influences on Party Loyalty." *American Political Science Review* 53 (September) 757–76

Maccoby, E. E., R. E. Matthews, and A S Morton. 1954 "Youth and Political Change." *Public Opinion Quarterly* 18 (Spring)· 23–39

McConnell, T. R., and P Heist 1962 "The Diverse College Student Population." Pp 225–52 in *The American College,* edited by N Sanford New York· Wiley

Matthews, D. R , and J W. Prothro 1964 "Southern Images of Political Parties." *Journal of Politics* 26 (February): 82–111.

Nie, N. H., S Verba, and J R. Petrocik. 1976. *The Changing American Voter.* Cambridge: Harvard University Press.

Niemi, R. G 1973. "Collecting Information about the Family: A Problem in Survey

Methodology." Pp. 476–90 in *Socialization to Politics*, edited by J. Dennis. New York: Wiley.
Pomper, G. 1975. *Voter's Choice*. New York: Dodd, Mead.
Rose, R. 1974. *Electoral Behavior: A Comparative Handbook*. New York: Free Press.
Spady, W. G. 1967. "Educational Mobility and Access: Growth and Paradoxes." *American Journal of Sociology* 73 (November): 273–86.
Stouffer, S. A. 1955. *Communism, Conformity and Civil Liberties*. Garden City, N.Y.: Doubleday.
Thompson, K. H. 1971. "Upward Social Mobility and Political Orientation: A Reevaluation of the Evidence." *American Sociological Review* 36 (April): 223–234.
Trow, M. 1961. "The Second Transformation of American Secondary Education." *International Journal of Comparative Sociology* 2 (March): 146–66.
Webster, H., M. B. Freedman, and P. Heist. 1962. "Personality Changes in College Students." Pp. 811–46 in *The American College*, edited by N. Sanford. New York: Wiley.

Language, Structure, and Measurement: Structuralist Semiotics and Sociology[1]

Charles C. Lemert
Southern Illinois University at Carbondale

Recent literature documenting the necessary relationship among the concept of social structures, the practice of measurement, and the nature of sociological language is reviewed. Structuralist semiotics (with special reference to Saussure) is used as a resource for examining the deficiencies of both formalist and interpretativist positions on these relationships. Structural explanation is described as an alternative. Umberto Eco's idea of a semantic encyclopedia is used as the basis of an illustration which is set forth in contrast to a formalist (Blau) and an interpretativist (Bittner/Garfinkel) theory of social structures. The argument rests upon the assumption that, because measurement is the encodement of what sociologists call observations, structuralist semiotics can be employed to provide an explanation for the language of measurement that depends on a semantics of social structures as such.

In sociology there is a necessary bond between language, structure, and measurement. This is well illustrated by a text from Coleman:

> The extraordinary complexity of structures of relations between people makes the task of measuring them a difficult one . . [T]he conclusion was reached [Coleman 1964, chap. 2] that quantitative measurement, to prove theoretically useful, had better be carried out in direct conjunction with the theory within which it is intended to be used. On an intuitive level, it comes down to the simple fact that there is no use developing a concept of "segregation" or "stratification" unless there are some definite causal relations in which the concept plays a part; and once such causal relations are made explicit, they dictate many of the decisions about how to construct a measure of the concept. [Coleman 1964, pp. 430–31]

What Coleman is saying is that a social structure is, above all else, a concept in the sociologist's technical vocabulary, not a self-evident real-world object. The causal relations in which the concept participates are specific to science, not "reality." Since, in Coleman's view, measures are largely defined by these concepts, measurement (like structure) is an aspect of sociological language and not primarily the literal representation of an existing

[1] I am grateful to the anonymous reviewers. Their very thoughtful critical texts are now part of this article's context.

social object. If measures are derived from theoretical concepts, measurement values are necessarily determined more by sociological language than by presumably real empirical facts. In other words, social structures exist in the theoretical language which establishes the conditions of their measurement.

This view is widely held especially by those who are best known for their interest in the formal measurement of social structures. It is essentially the same as Blalock's view of indirect measurement and auxiliary theory: he argues that measurement theory must be included in the abstract theory of structural relations (Blalock 1968*a*). Like Stinchcombe (1968, pp. 28–29), Blalock makes the structural variable a theoretical definition, not an empirical sign (see Lemert, in press *a*). And Parsons's analytic realism methodology hinges entirely upon the establishment of structures within a preempirical theoretical lexicon: "A structure is not an ontological entity but is strictly relative to the investigatory purpose and perspective" (Parsons 1970, p. 35; cf. Blau 1974).

Recently this discussion has been joined by a group of sociologists whose interests are otherwise quite different from those of Coleman, Blalock, Parsons, and Blau. Since the appearance of Cicourel's *Method and Measurement* (1964), ethnomethodologists have explicitly concerned themselves with the relationship between language and the sociological measurement of structures (e.g., Garfinkel 1967*b*; Garfinkel and Sacks 1970; Bittner 1974; Cicourel 1964, 1974). Ethnomethodology has claimed that measurement becomes a universal and necessary feature of both professional and practical sociological talk when measurement is understood as that social process whereby encoded norms are employed to make sense of the social world.

> We call attention to the phenomenon that formal structures are available as accounts of professional sociology where they are recognized by professionals and claimed by them as professional sociology's singular achievement. These accounts of formal structures are done via sociologists' mastery of natural language, . . . This assures to professional sociologists' accounts of formal structures its [*sic*] character as a phenomenon for ethnomethodology's interest, not different from any other members' phenomenon where the mastery of natural language is similarly involved. Ethnomethodological studies of formal structures are directed to the study of such phenomena, seeking to describe members' accounts of formal structures wherever and by whomever they are done, while abstaining from all judgments of their adequacy, value, importance, necessity, practicality, success or consequentiality. We refer to this procedural policy as "ethnomethodological indifference." [Garfinkel and Sacks 1970, p. 345]

It is true, of course, that ethnomethodologists use nonstandard meanings for measurement and social structure. However, it is clear that by making language a primary theoretical topic they are considering the same problematic as Coleman and others.

It is, therefore, the present catholicity of the language/structure/measurement problem which occasions its reconsideration in this paper. I will argue, however, that neither of the two general points of view mentioned above is satisfactory. The first, of which Coleman is typical, may be called *formalism*. Here "structure" usually refers to macro structures. Language is that which must be purified in order to make measurement possible. The second approach, of which ethnomethodology is typical, is commonly identified as hermeneutical or interpretative sociology.[2] Here structure is a weak term, often referring to an aspect of the individual's interpretative repertoire, such as an actor's ability to recognize and organize "socially acceptable behavior" (Cicourel 1974, p. 11) or "joint lines of action" (Blumer 1969, p. 17). However, the weakness of this usage is somewhat repaired by a strong analysis of the way in which language determines all measurement procedures.

Both approaches see the measurement of structures as problematic. However, the formalists see technical sociological language as that which can overcome measurement problems, while interpretativists see language itself as the problem which dooms all measurement to endless self-examination. This is the difference between an occasionally humble heroism and a faintly optimistic fatalism. Accordingly, while considering the same general issue from different perspectives, each emphasizes a different aspect of the problem. The formalists, though naive as to the complexities that prevent language from achieving any pure formal state, have properly emphasized the fact that the concept *structure* must, in sociology, get at macrosocial structures in a clear and precise way. The interpretativists (especially ethnomethodologists), though seemingly uninterested in the complexity of structures in a macro sense, have properly insisted on the relativity that language necessarily imposes on any scientific discourse.

The problem with which I will deal is: How can sociology measure complex social structures while retaining a proper respect for the contextual relativity that language imposes? Because I think that formalism and interpretativism give unsatisfactory answers, this article will seek a solution in a third alternative, the structuralist social theory that derives from Ferdinand de Saussure.

STRUCTURALIST SEMIOTICS AND SOCIAL THEORY

Semiotics is the study of signs and their meaning. Semiotics is related to linguistics, but language is only one of many sign systems. In addition to such sign systems as Morse code, highway signs, and fashion codes, one may

[2] The classification used here is roughly comparable to others in use (e.g., Wagner 1963; Douglas 1970; Habermas 1971; cf. Lemert, in press *a*).

properly include measurement codes employed by sociologists as a topic of semiological research. Structuralist semiotics is a prominent tradition of semiological research which (for reasons given below) is of special interest to sociologists. Ferdinand de Saussure—a Swiss linguist who died in 1913—is generally considered to be the founder of structuralist linguistics (Chomsky 1972; Jakobson 1974; Culler 1976; Lepschy 1970; Reid 1974; Foucault 1972) and of structuralist semiotics (Barthes 1970; Eco 1976).

The precise importance of structuralist semiotics to sociology is suggested by the fact that Saussurian semiotics has led to important accomplishments in social theory in Europe. The work done by Lévi-Strauss in anthropology, by Jacques Lacan in psychoanalysis, by Michel Foucault in social history, and by Umberto Eco in semiotics has been stimulated (directly or indirectly) by Saussure. There is, however, a more fundamental warrant for this consideration of structuralist semiotics. Though Saussure himself never claimed to be anything more than a linguist, his principal text (*Course in General Linguistics* [1959])[3] can be read as a general social theory relevant to the problems of interest in this article.[4] In particular, the *Course* demonstrates that the scientific observation and measurement of language facts are necessarily dependent upon a systematic theory of language as a complex social structure.

Saussure worked out this theory with reference to the famous distinction between *la langue* (language) and *la parole* (speaking). The latter refers to the observed data of linguistic research, the actual speech of language users. However, left to itself, *la parole* overwhelms the linguist with a large number of disorderly facts. When one takes into account the variety of words, phonemes, and tonal emphases actually used by even the most elementary of speakers (strangers, children), it is apparent that speaking involves a virtually unlimited number of events. Science is inundated. Order is difficult to achieve.

Saussure overcame this superabundance of fact and established linguistics as an orderly science by means of the concept *la langue,* which he viewed as the social system of conventions from which *la parole* draws its elements. "Language [*la langue*] is a well-defined object in the heterogenous mass of

[3] The *Course* was not actually written by Saussure. It is a compilation based upon students' notes of lectures given at Geneva in 1907, 1908–9, 1910–11

[4] The fact that Saussure placed linguistics within semiology (1959, pp. 15–17) is the main evidence that he too saw his linguistics as a general social theory. Likewise, his theory of the relationship between the individual and society (1959, pp. 13–15) and his view that linguistics is subsumable under all value-specific sciences (1959, pp. 79–81) support this interpretation The social theoretical implications of his work were developed by others (including Lacan, Foucault, Derrida, Lévi-Strauss, and Eco). Saussure is used here in order that the theoretical value of structural semiotics can be seen in its simplest form and in its historical context Saussure also makes it clear that "structuralism" is *au fond* a semiotic discipline, not a metaphysics of structures.

speech facts. . . . It is the social side of speech, outside the individual who can never create or modify it by himself; it exists only by virtue of a sort of contract signed by members of the community" (Saussure 1959, p. 14).[5] Though, for Saussure, linguistic observation must take its units from speaking (*la parole*), linguistic analysis can define these units as facts only by locating them in *la langue*. With Saussure linguistics, in effect, becomes an implicitly sociological exercise insofar as facts are explained by a social convention, by a community's most elemental social structure: its language.

The *la langue/la parole* distinction is supplemented by another crucial dichotomy, *paradigm/syntagm*.[6] Roughly described, Saussure's paradigms are the content of *la langue;* that is to say, the series of formal structures (verb forms, word variations, grammatical rules, phonetic signals, etc.) from which the speaker selects the units that are formed into actual speech chains (syntagms) which are the substance of speaking (*la parole*). The syntagms are what the listener is aware of (as is, of course, the linguist); the paradigms always remain absent to observation. The latter are the structures which combine in an extremely complex manner to form the structured social convention, *la langue*.

Thus, for Saussure, linguistic measurement is a description of the place of observed units of speaking (*la parole*) in a social structure (*la langue*). It is here that one begins to see the value of Saussure's alternative to formalist and interpretativist sociologies. His notion of *la langue* resembles ethnomethodology in that it pictures social structures as conventions and thus as negotiated. However, unlike ethnomethodologists (and interpretativists in general) Saussure shares the formalist interest in providing formal descriptions of macrosocial structures. With respect to language, *la langue* is a formal macro structure insofar as language structures are not considered immediately susceptible to manipulation by speaking individuals (Saussure 1959, p. 74). In short, Saussure's position retains the social naturalism that interpretative sociology has always considered a fatal omission of formalism, but does not reduce the formal study of social structures to a metaphoric microsociology of intentions (the fatal flaw of interpretativism, in formalism's eyes).

[5] It is almost certain that Saussure, who lectured in Paris at the Ecole des Hautes Etudes from 1881 to 1891, was influenced by Durkheim In this passage one sees the probable influence of Durkheim's *conscience collective*. Their views of social fact are, also, nearly identical (compare Saussure 1959, p 77, with Durkheim 1938, p 13). The crucial difference is that Durkheim provided no technical concept comparable to Saussure's *la parole* and therefore did not explicitly explain the technical basis for the measurement and analysis of the relationship of individuals' motives to social structures. The advantage of Saussure's implicit social psychology will be illustrated below.

[6] The paradigm/syntagm distinction is roughly comparable to more familiar dichotomies such as code/message, competence/performance, and metaphor/metonymy (see Barthes 1970, pp. 18–20).

STRUCTURALIST SEMIOTICS AND MEASUREMENT

The value of Saussure to sociology rests upon our ability to think of sociological measurement as a language-specific semiotic process able to generate formal descriptions of macrostructures. In order to develop this possibility, it is necessary to begin with a discussion of Saussure's theory of the sign (the unit of linguistics) and its value (the sign's measurability). This requires an examination of two additional Saussurian ideas: (1) the arbitrariness of the sign, and (2) the differential character of linguistic value.

Arbitrariness

"The linguistic sign unites, not a thing and a name, but a concept and a sound image" (Saussure 1959, p. 66). This is Saussure's simplest definition of the sign, the unit of structural linguistic research.

The sign is nearly as important for what it is not as for what it is. It is not a way of naming objects in the real world. The term "cow" is not the name for a class of milk-producing bovine creatures. For Saussure, it is instead a psychological construct by which the speaker unites his or her acoustic impressions of certain phonemes with the concept *cow*. The "real" cows people take their children to gawk at have no necessary connection with the linguistic sign. Saussure, at this point, is both antiempiricist and antinaturalist. Methodologically, his unit of analysis does not pretend to present properties of the external world. Philosophically, he does not believe that the natural object-world is of primary interest.

This is not to say, however, that Saussure is a psychological reductionist or a methodological individualist.[7] The sign is not arbitrary in the sense of allowing any speaker to use any combination of sounds he chooses as a signifier. "The signifier, though to all appearances freely chosen with respect to the idea that it represents, is fixed, not free, with respect to the linguistic community that uses it" (Saussure 1959, p. 71). The sign is a social event in that speakers must use the linguistic conventions established

[7] This is a problem in Saussure. On the one hand it can be said that he avoids psychological reductionism by founding the acoustic impression in sound. The impression thus derives from the material fact of speaking itself, while simultaneously being associated with a concept. There could be no sign if one were to consider concept, acoustic impression, and sound in isolation from each other. At the same time, there is a marked psychologism here. This is one of the points at which Saussure's concept *sign* has come under attack Derrida (1976), while recognizing the importance of Saussure's concept, argues that it retains an essentially metaphysical taint. Eco (1976) claims that it lacks the precision necessary to explain its role in semiosis and that it is better replaced by the concept *sign-function*. However, since the present discussion concerns other, more general, features of Saussure's theory, these criticisms are not of direct importance here, though they would be in a complete development of Saussure's social theory. That Saussure is not a psychological reductionist can be readily seen by comparing his views with, for example, Homans's (1967) explicit methodological individualism.

by a community of speakers. These conventions include phonic and lexical items and syntactical rules. The sign is the unit of the speech chain (the syntagm) which draws its substance from the paradigms of a given language (*la langue*).

Behind Saussure's theory of the sign is an assumption of considerable importance. Language is not directly the product of subjective intentions. Even the sign articulated by an individual is formed with reference to social conventions. Thus, Saussure's idea of arbitrariness has two dimensions: (1) Signs are arbitrary insofar as they are social conventions. (2) They are arbitrary insofar as they are composed of elements (a concept and a sound image) that have no necessary or natural connection to each other. Neither of these ideas is entirely foreign to sociology.

The first dimension (that units of social behavior are arbitrary because of their origin in societal conventions) has been long accepted by various interpretative theories. In this view the observed units of social behavior ("indexical expressions," "behavioristic specimens," "joint acts," "frames," etc.) are the products of social negotiations. Thus, since they are not given, they are not readily formalizable. Harold Garfinkel's well-known definition of ethnomethodology is a convenient illustration: "[T]he term 'ethnomethodology' . . . refer[s] to the investigation of the rational properties of indexical expressions and other practical actions as contingent accomplishments of organized artful practices of everyday life" (1967*b*, p. 11). In other words, the contingent nature of social behavior limits measurement to indices of the actor's intended meaning. Though nuance is necessarily overlooked here this is, roughly, the position of all interpretative sociologies from symbolic interactionism, through Goffman, phenomenology, and ethnomethodology to Habermas's "depth-hermeneutics" (Habermas 1970*a*).[8]

The formalists, on the other hand, beg the question of the contingency of social behavior with respect to hidden cultural meanings in order to clear the way for the designation of formal units of analysis which are presumed to function as measures of properties of a real world. Though it is seldom recognized, this position is arbitrary insofar as the theoretical variable is not a natural fact but the theorist's definition of an indicator. Take, for example, Stinchcombe's description of a variable as ". . . a concept which can have various values, and which is defined in such a way that one can tell by means of observations which value it has in a particular occurrence" (1968, pp. 28–29). The key words are "defined in such a way." Formal

[8] Though I take ethnomethodology as a representative of interpretativism, this is not to say that differences within this larger perspective are negligible. Ethnomethodology is of primary interest here because it has made explicit prior interpretativists' implicit discussions of the language/measurement/structure problem. For discussions of the place of ethnomethodology in this tradition see Attewell (1974) and Lemert (in press *b*). For discussions of Habermas's place in interpretativism see Habermas (1970*a*, 1970*b*, 1971) and Lemert and Gillan (1977).

variable definitions are arbitrary in the sense that they are determined by the internal logical or mathematical requirements of the theoretical system, not their representative exactness. "The values of the structural parameters, which provide us with a mathematical explanation of the empirical relationships, cannot be determined from the data alone" (Blalock 1970, p. 284). Though in another respect (to be discussed) this position contradicts itself, here formalism is surprisingly similar to Saussurian semiotics in that its unit of analysis (the formal variable) does not name real-world properties.

Where Saussure and both sociological approaches are in accord is on the idea that, in the social sciences, an absolutistic theory of measurement cannot yield theoretically useful results. Unless one is willing to settle for a naive empiricism such as operationalism, social research must admit some degree of arbitrariness into its theory-construction strategies in order to deal with a complex social world that does not give up its facts to a direct measurement. Interpretativism attacks this problem by emphasizing the arbitrariness by which the observed units are constructed, while formalism takes the alternative of concentrating on the arbitrariness of the theoretical constructs (variables) by which facts are explained. Unfortunately, neither of these tactics is satisfactory by itself. The former has not yet shown how the arbitrariness of its units can be explained formally with reference to structures. The latter (formalism) has not been able to deal with the contradiction involved in describing structures by means of arbitrary theoretical variables derived apart from the direct measurement of given (and absolutistic) units.

The Saussurian alternative is not, however, merely a convenient historical reference for a synthesis of ethnomethodology and formalism. Saussurianism, while touching on similar issues, leads sociology to a complete reformulation of the relationship among structures, measures, and facts. The value of such a reformulation depends on the assumption that social behavior follows laws similar to those of language behavior. The development of this assumption for sociology derives from the second major part of Saussure's theory of signs, the differential character of linguistic value.

Differential Values

One consequence of the arbitrariness of the sign is that it is present to observation as a unit of speaking, but cannot be explained scientifically without reference to the structure of its language (*la langue*), which is absent and not directly observable:

> . . . [W]e must first be convinced that the concrete entities of language are not directly accessible. If we try to grasp them, we come into contact with the true facts. Starting from there, we can set up the classifications that linguistics needs for arranging all the facts at its disposal. On the other

> hand, to base the classifications on anything except concrete entities—to say, for example, that the parts of speech are the constituents of language simply because they correspond to categories of logic—is to forget that there are no linguistic facts apart from the phonic substance cut into significant elements. [Saussure 1959, p. 110]

In other words, the articulated sign is the unit of observation, but the "true facts" of language are found in the paradigms making up the structure of *la langue*. The former is present, the latter absent (Saussure 1959, p. 123).

The effect of this interpretation is that the linguistic value of the term "cow" is only partially found in its signifying function (that is, in the linkage between the phonic signifier "cow" and the signified concept *cow*). Value depends also on the relations between any enunciated term and (*a*) the other elements with which it is combined in an actual speech chain, and (*b*) those paradigmatic elements which were omitted from the utterance. Therefore, the value of "cow" is based upon its differences from other linguistic units. When "cow" appears in the sentence "I milked the cow," its value is different from that in the sentence "The cow milked me." The different combination in the latter case does nothing to the sign "cow" itself, but it dramatically alters the value of the term by connoting a highly unusual cow, such as one that might appear in a dream or a children's book. Even when "The cow milked me" connotes "The cow swindled me," the cow is transformed into a figurative creature to which human intentions are imputed. In addition, the value of "cow" in an utterance depends on the fact that the term "cow" appears, while innumerable possible alternatives have been omitted: such as, "I milked the *goat* (or *kitten, welfare office*)" and "The *gynecologist* (or *my husband, the nurse*) milked me."

For the present, the most intriguing of these differential relations is the latter, that between the present enunciated term and its absent paradigmatic possibilities. In this connection, Saussure attempted to relate linguistics to other social sciences that concern themselves with value (economics, sociology, anthropology): "Instead of pre-existing ideas then, we find . . . *values* emanating from the system. When they are said to correspond to concepts, it is understood that the concepts are purely differential and defined not by their positive content but negatively by their relations with the other terms of the system. Their most precise characteristic is in being what the others are not" (Saussure 1959, p. 117).

The argument now rests on the plausibility of Saussure's generalization. Is the value of social action also determined by a differential system? Is its character described best differentially, that is, as being what other unactuated possible actions are not?

In sociology the debate boils down to two basic positions. One (the formalist) assumes—in contradiction to the arbitrariness of its variables—that

the value of a given behavior (or pattern of behaviors) is a self-evident property of a social world. It is therefore observable (though usually by indirect measures). "The reason for having theories of social phenomena is to explain the pattern in observations of the world" (Stinchcombe 1968, p. 5). For such a statement to hold, it must be assumed that the actions observed are in no final sense mysterious. This permits the classification of observed action values with respect to variable categories. In other words, the scientific values (those which constitute the measurable features of variables) are taken as accurate representations of human values. The "meanings" of behavior patterns in a social group are described by the empirical measures of that behavior. This amounts to the assumption that the value of an action is inherent as a property of the observed actor. There is no insurmountable problem of interpretation. The actor does not hide any significant values from the observer.

There is, as I have already suggested, a contradiction here. It is important to distinguish between the measured value of observed behaviors, which is an aspect of measurement, and the relationship between indicators and variables, which compose a theoretical system (see Blalock's [1968*a*, 1970] discussions of auxiliary theory). The former relationship (observations to measures) concerns values and is not considered arbitrary; the latter (indicators to variables) concerns the units of analysis and is arbitrary. In a theorist such as Blalock this distinction points to the often overlooked fact that theory constructionism is anything but empiricist. On the contrary, it places so much emphasis on the problems of the theoretical system itself that empirical values become a secondary consideration. Thus is created the contradiction between the relativism of the theoretical system and the crypto-absolutism guiding empirical research. This contradiction is the source of the flaw in formalism, which has relied upon formal language as the necessary precondition for replication studies. Replication is understood as the procedural requirement for the accumulation of knowledge—that is, for the maturing of sociology as a science. However, this procedure is undermined by the fact that replication is sought by means of arbitrarily derived abstract theories which rest on the flimsy justification of the scientists' own theoretical judgment. Were formalism to take the other course of gaining formality (and the possibility of replication) in the absolutism of given social facts (where theory is merely the name for observations), it would be trapped in the naive operationalism which virtually every major contemporary formalist desires to avoid (Parsons 1937, p. 72; Blalock 1968*a*, pp. 7–9; Coleman 1964, p. 63; cf. Kaplan 1964, pp. 39–42). The significance of this flaw is found in the fact that formalism, in spite of its good intentions, has not come anywhere near to becoming a formal science based on replication studies (Blalock 1973).

Interpretativism, on the other hand, has frankly conceded that one cannot easily measure the value of behaviors because value is a function of subjective meaning which is not directly observable. Blumer's position is typical: "The position of symbolic interactionism . . . is that the meanings that things have for human beings are central in their own right. To ignore the meaning of the things toward which people act is seen as falsifying the behavior under study" (1969, p. 3). The value of a unit of behavior is subsumed under its meaning to the actor and, therefore, cannot be easily counted.

What is clear is that neither of the two sociological views of value is as complete as Saussure's. Both assume that the value of a social act is representative of a natural property of social life and therefore can be represented as a value in sociological propositions. In the formalist view, value is inherent in the behavior itself. The question of its origin is ignored so that attention can be given to its formal (often mathematical) representation within a theoretical system. In the interpretative position, the value of a social act is created by the actor. Though this makes representation within theory more difficult it is clear that interpretativists are not reluctant to represent meanings in verbal theories. In both cases, value is seen as a positive feature of the social act (given in one, created in the other). Both positions are, therefore, quite different from Saussure's, wherein value is a purely negative entity, determined entirely by the differential system. In other words, for Saussure the value of a unit of action is dependent upon the social structures of which that unit is a part; the individual cannot be explained without the social. Measurement can proceed only by the description of the values formed in social structures.

The Saussurian alternative is summarized in two principles, both of which are applicable to sociological measurement: (1) that neither the units of observation nor the measured facts derived therefrom are direct representatives of a real empirical world; and (2) that, nonetheless, measurement is possible by the derivation of negative values from a differential structure. Saussure resolved the seeming contradiction here by his strongly sociological (and Durkheimian) concept of *la langue*. If the signs which become units of linguistic measurement are not representatives of reality (are not names for things), then what is being measured? Since it is the society that creates *la langue,* then to describe linguistic units (signs) with reference to that convention is also to examine a real social structure.

Obviously, this works very much to the convenience of one who is studying language. But the question that remains is: Can such a program be generalized to extralinguistic phenomena? In other words, is structural explanation possible in sociology?

STRUCTURAL EXPLANATION AND SOCIOLOGY

Science measures in order to explain. Measurement (however it is defined) can have no other scientific purpose. The usefulness to sociology of a structural semiotics of explanation depends on the answer to two questions: (1) Is the structural explanation derived from Saussure sufficiently original to improve upon practices currently used by sociologists? (2) Is it sufficiently explanatory to be sociological?

Originality

Why bother with Saussure if the same results could be achieved by an elaboration of the existing formalist or interpretative theories? The question arises because there are obvious similarities.

Saussurian semiotics is similar to formalism in that both are structural. When Coleman insists on lodging measurement within formal theory or when Homans and Blalock insist that explanatory theory is a system of "propositions that interrelate variables" (Blalock 1970, p. 274), they are using a kind of structural explanation. In formalism, the structure is a theoretical system of concepts and propositions within which values associated with empirical observations are interrelated. The empirical instance is always referred back to a structure, just as Saussure refers *la parole* back to *la langue*. The limitation of the formalist method is that it cannot explain the structures (deductive systems) it creates. Whereas a Saussurian approach would begin with the assumption that even scientific language is necessarily referrable back to social conventions, formalism must refuse this step because of the damage it would do to the precision of formal statements. One cannot easily maintain the generality of a formal statement if its source is the relativity of a community's conventional language. As a result, formalism buys precision at the high cost of empirical accuracy. "The cost of greater simplicity is the sacrifice of realism" (Blalock 1970, p. 296).

On the other hand, if Saussurian thought corrects formalism by introducing language, how does it differ from those interpretative approaches, such as ethnomethodology, which base themselves upon the centrality of language to sociological measurement? When Garfinkel defines ethnomethodology as the "investigation of indexical expressions" he necessarily makes sociology responsible to language (Garfinkel and Sacks 1970, p. 338; cf. Sacks, in Sudnow 1972, p. 31; Cicourel 1974, p. 11; cf. Lemert, in press *b*). Moreover, he does so in a fashion that appears to permit a Saussurian analysis of the indexicality of everyday life phenomena in terms of a description of the differential elements (paradigms) constituting language itself. Garfinkel's description of formal structures in accounting discourses (Garfinkel and Sacks 1970, pp. 355–56) and Cicourel's lists of interpretive procedures (1974, pp. 51–58) seem to take this form. However, what prevents these

major ethnomethodological positions from achieving the explanatory scope of Saussure is (not incidentally) the same thing that requires them to use a diminished and social-psychological concept of social structure. The interpretative position has arrived at language through the back door. It begins, not with language (*la langue*), but with a dominant feature of everyday life: the motivatedness and constructedness of social behavior as seen in speaking (*la parole*). "Whenever and by whomever practical sociological discourse is done, it seeks to remedy the indexical properties of practical discourses . . ." (Garfinkel and Sacks 1970, p. 339). Sociological reasoning—though bound to language and, therefore, given as a property of all social creatures—is a remedial enterprise. That is, it remains a hermeneutical practice, not an explanatory discipline. Ethnomethodology shares the fate of all hermeneutical methods: an inescapable bondage to the intersubjective history in which social reality is assumed to be constructed (see Habermas 1971, 1973). Thus, whereas Saussure explains individual actions (*la parole*) by reference to a social structure (*la langue*), interpretativism reverses this order. When social structures (including *la langue*) are viewed as the product of intersubjective scheming, one necessarily undermines the theoretical autonomy of complex social structures. The latter are virtually lost to precise theoretical work because of the overwhelming importance given to intersubjective motives.

Applicability

To pass from Saussure's originality to his applicability is to question the usefulness of structural explanation in sociology. Since "structuralism" is in direct and intentional conflict with causal explanation, it challenges the only strong theory of explanation sociology has ever used.[9]

Structural linguistics cannot ask why (what caused?) the speaker to utter (select) certain phonemes and words. To do so is to turn away from *la langue* to the question of a speaker's motivations and, thereby, to move explanation outside linguistics proper to ethics or psychology.

Structural explanation rests upon the description of the systems of differences from which utterances are selected. What permits explanation is the

[9] To be more precise one must say that structuralism challenges all historicist methods (see Foucault 1972; Lévi-Strauss 1966; Althusser 1970) Causal reasoning is here taken as the sociologically most important type of historicist reasoning. Stinchcombe (1968) has suggested that all three basic sociological strategies (functionalist, historicist, and demographic) are historicist (cf. Lemert, in press *a*) Even though it is the formalists who have sponsored causal explanation (e.g., Homans 1967; Stinchcombe 1968; Blalock 1968*b*), those interpretativists who have attempted more than descriptive studies have fallen back on a type of causal reasoning (e.g., Blumer 1969) The argument that predictiveness, not causality, is the standard of explanation in sociology (Gibbs 1972) does not threaten this position, since predictiveness is also a variant of historicist reasoning in the broad sense.

idea that the scientifically interesting facts are those of the social structure, *la langue*, not *la parole*. Thus, structural explanation does not seek the origins of linguistic behavior because all of the usable elements of speaking behavior are assumed to have been defined by the social community.

Whether this can work in sociology depends on the relationship between the particular structure language and social structures in general. If the social formation of language follows unique and nongeneral rules, structural explanation does not apply. If, on the other hand, social structures generally can be viewed as a set of absent, predefined, differential elements, structural analysis may well be a useful form of sociological explanation.[10]

What exactly is a social structure such as a given stratification system? In the formalist approach, it is a system of concepts imputed by the sociologist to a set of data that is assumed to reflect an unequal effect of social norms and material conditions on social behavior. In the interpretative approach, such a social structure is the sociologist's second-order construct for a primary social construction. Both are causal explanations. In the formalist case, one assumes that a class of individuals is "acted upon" by structured causal forces which the sociologist formally specifies in terms of a system of propositions. In the interpretative case, social individuals are viewed as creators of a stratified social structure which reflects their understanding of a meaningful life. Stated generally: in the former case, stratification causes behavior; in the later, interpretative behavior causes stratification. Neither can escape the stern linearity of causal reasoning wherein social structures and individual, intersubjective behaviors must be conceived as belonging to two separate orders of reality. More to the point, both partake of the assumption of causal analysis that measures of structures and behaviors must be stated in terms of positive values. In effect, this assumption is linked to a broader philosophical belief that social facts (structures and behaviors) have a positive existence.

But do they? In what sense is the "existence" of a stratification system positive? Where does "upper middle classness" exist except in the language of sociology? For people "living" within a given stratification system, that structure exists not as an evident positive fact, but as a concrete set of limitations and possibilities. A structure "functions" as a series of differential choices (made, makable, or only imaginable) presented by means of

[10] Though the statement sounds extreme it is supportable. Two issues are involved: First, one must note that causality is very much bound to classical metaphysical principles that require all thought to derive from some central, first principle (see Derrida 1976; Foucault 1972). Therefore, to the extent that social science wishes to be empirical it must, at least, explain how it deals with this corrupting influence. Second, to say that sociological explanation is structural is not to say that causal or historicist reasoning has no place in sociology (e.g., Saussure's distinction between diachronic ["historical"] and synchronic ["structural"] linguistics). One might consider the same sort of division of sociological labor. Causal and historical studies would belong to a diachronic sociology that would study the origins of and changes in social structures.

various normative and material paradigms containing differential possibilities and prohibitions with respect to self-esteem, manners, income, housing, standard of living, etc. Nor is it self-evident that "behavior" is positive. In the structuralist view, behavior is theoretically negative in the sense that a given behavior is an activated formal possibility, not a positive intentional content generated by an actor. It is the selection of a form from structured paradigms. The forms selected are not free-standing (positive) objects, since they are determined only by their difference from other forms of normal behavior, which exist only in the social structures.

By contrast, linearity and positivity are necessary and concomitant features of causal reasoning. The attempt to explain the formal properties of causal relations requires positive theoretical objects, separated linearly from each other and, in principle, "possessing" different "forces." The result is ultimately fatal to any nonabstract explanation of complex structures because one has no alternative but to reduce the complexity of structures to insufficiently subtle formal propositions. Saussure provides a way to avoid this trap.

Described in terms of its semiotic[11] (and most basic) function, measurement is the application of a standard conventionalized code to a series of observations. Causal measurement, therefore, seeks to generate positive facts that are taken as names for "observed" real-world properties. The names are given by presumably standard (formal or "natural") codes that are believed to be capable, in turn, of generating the positive values which are explained by a theoretical structure. Structural measurement, in contrast, seeks to isolate arbitrary, negative facts observed not in a presumptive real world, but in concrete events constituting articulated texts of social behavior, which are observable only after measurement. The measure, in this latter case, is not an abstract standard but a conventional code in the form of a system of differences (see Eco 1976, chap. 2). Structural measurement, therefore, seeks to locate articulated behaviors in a particular semantic field.

In causal reasoning, positive entities are measured; in structural reasoning, differences are measured. In the former, science imagines linear spaces in which objects move with respect to paths. In the latter, science imagines temporary semantic fields that locate the positions of encoded signs of practical behaviors (that is, tokens).[12] In the one case, measurement is the act

[11] Semiotics, the science of signs and signification, studies all linguistic and extralinguistic (fashion codes, animal communication, medical diagnosis, etc) systems of signification (see Barthes 1970; Saussure 1959; Eco 1976, pp. 1–29). Below, it will be shown that scientific measurement and explanation constitute a semiotic process in that they involve the encoding of empirical signs.

[12] The concept *token* and its correlate *type* are necessary to overcome the linguistic and sociological realisms which assume that empirical signs represent real-world objects (see Eco 1976, pp. 178–88, 217–60).

of projecting a standard code along an axis of observations; in the other, measurement refers observed distinctions back to the social structures by which the distinctive units are observed. Causal measurement, therefore, requires abstract context-free codes that deny even their own conventional nature.[13] Structural measurement, in contrast, requires temporary scientific standards predefined by the same social structures they measure. In other words, a causal theory of measurement (such as, for example, Blalock's) is a theory of observational strategies, while a structural theory of measurement is a theory of intercode (interstructure) relationships. Or: the former is a theory of scientific procedures, the latter of social structures.

Therefore, since extralinguistic social structures are conceivable in Saussure's sense and since Saussure does offer the outlines of a theory of explanation, structural reasoning could assume a central place in sociological work. Sociology, thereby, would be concerned primarily with the analytic description of the paradigmatic differences and negative values that constitute social structures.

STRUCTURAL EXPLANATION AS SEMANTIC ENCYCLOPEDIA

Finally, one must ask whether structural explanation can produce general scientific statements. The proper criticism of interpretative sociology is that it can generate only second-order descriptive texts which are not sufficiently more general than the first-order texts (everyday behavioral events) it studies. It is too concrete. Formalism seeks to correct this deficiency by translating verbal texts into formal, mathematizable lexicons which are then manipulated into general propositions. The proper criticism of formalism is that it gains general explanations while losing the richness (and "reality") of descriptive detail. It is too abstractly general.

Structuralism is an alternative that retains the concreteness of interpretativism and the generality of formalism. This accomplishment is based upon a frank reappraisal of explanation. Explanation is understood—from the structuralist point of view—as a semantic process that generates a certain type of statement: namely, one that *meaningfully* encodes already encoded measurement values. Explanation is a semantics of sociological signs. If, then, explanation is the production of sociologically *meaningful* ("*valid*") *statements*,[14] the problem of generality in explanation must be reformulated with reference to a semantic theory.

[13] The sociological problem of context-free abstractions is discussed in Cicourel (1974) and Habermas (1970*b*, 1973); cf. Eco (1976, pp. 96–105).

[14] Quotation marks and italics are used in this section as diacritical marks with intentional, technical purposes. Quotation marks denote terms used in a more or less sociologically normal manner. Terms thus marked are assumed, however, to be either problematic or in need of further explanation which cannot be given here. Italics mark terms used in a semiotically technical fashion specific to this text. Thus, *meaningful* here refers not to a commonsense qualification of *statements* but to the task of a semiotic semantics

It is necessary, at this point, to shift from Saussure's classical semiotics to more recent work in structuralist semiotics. The most cogent and pertinent to sociology is that of Umberto Eco, who has distinguished between dictionary and encyclopedic semantic theories (1976, pp. 98–100). In these terms, formalism provides a dictionary-like generality: that is, its statements are abstract, context-free definitions. (Interpretative explanations are undercoded descriptive texts; that is, context-absorbed commentaries.) Structural explanations are, in contrast, encyclopedic texts having the following features: Though concrete, they are more general than interpretative descriptions. They are not, however, abstract like formalist dictionaries. An encyclopedia provides contexts. Their mathematizability is potential, even probable, but not necessary.

Since it is impossible to develop all of the logical details of this alternative, it is best to explain encyclopedicity by means of an illustration, presented in contrast to both formalism and interpretativism.

The structure of formal organizations is an excellent empirical example. When sociologists discuss *formal structures* in this setting they must consider a general, macro feature of social organization. At the same time, formal structures in organizations cannot be isolated from their impact on such concrete behavioral problems as informal structures, personality needs of members, and decision-making processes. Given the broad consensus on this point (Haire 1975, pp. 4–12; Gouldner 1954, p. 27; Blau and Schoenherr 1971, pp. 4–12; Merton 1957, pp. 195–206; Bittner 1974, p. 69), one would expect sociological explanations of formal structures to revolve around the incongruity between general structures and concrete individuals. However, they do not. Sociologists usually establish a division of labor between the study of formal and informal organizational features (e.g., Blau and Schoenherr 1971, pp. 4–12; Blau 1970). While studying the one, they relegate the other to the background. Thus, a formalist such as Blau (1970) studies formal structures, dismissing for the occasion the question of individuals' intentions. Conversely, an interpretativist such as Bittner (1974) considers practical negotiations of formality in organizations while ignoring formal structures as such. This is not merely a matter of personal preference. It results from different measurement tactics. Codes used determine what can or can not be measured. I will illustrate this problem before demonstrating the advantage of an encyclopedic explanation.

---

to generate ***meaning*** in ***statements***. Likewise, "***valid***" is the sociologically encoded term which (in spite of its problems) serves the same semantic "function" as semioticians' ***meaningful***. Thus, ***xxxx*** denotes a technical usage explained in this text (e.g., ***measurement, type, token***); "xxxx" is a usage encoded elsewhere and not explained here (e.g., "knowledge"); "***xxxx***" represents a usage encoded here and elsewhere (e g., "***literature***").

Formal Dictionaries

Blau has written some of the most explicit attempts to formalize theories of social structure (Blau 1970, 1974; Blau and Schoenherr 1971; Blau and Duncan 1967). With reference to organizations, he has said: "Formal structures exhibit regularities that can be studied in their own right without investigating the motives of the individuals in the organizations" (1970, p. 203; cf. Blau and Schoenherr 1971, p. 6). The unasked question is: What are the risks taken when one makes such an arbitrary procedural distinction? Take, for example, one of Blau's general propositions about structures and one of its derivations: *"Structural differentiation in organizations enlarges the administrative component,* because the intensified problems of coordination and communication in differentiated structures demand administrative attention. . . ." "One derived proposition is that the large size of an organization indirectly raises the ratio of administrative personnel through the structural differentiation it generates" (1970, p. 213). It is, of course, possible that the relationship between organization size and the size of the administrative component is explained in part by members' sense of social distance from the organizational structure itself. In this case, it is likely that the increase in the administrative component is due, in part, to the managerial problem stemming from a decrease in members' sense of inclusion in the democratic process of shaping the organizational structure. Poorly integrated members must be managed, requiring a larger administrative component. Blau himself must have something of this sort in mind when he talks about "the intensified problems of coordination and communication in differentiated structures" (cf. Blau 1970, pp. 212–13). Elsewhere he states that the loss of democratic process in an overorganized society is "the challenge of the century" (1974, p. 633). But what is this problem of communication or possible malintegration, if it is not a problem of individual's motives which Blau, in this instance, has decided not to examine?

If, therefore, motives are likely to have an effect on the size of the administrative component, one must question the value of any "deductive theory of the formal structure of work organizations" (Blau 1970, p. 201) that, on methodological grounds, eliminates this topic from consideration. Here, again, one sees the contradiction within formalism wherein arbitrary (in Saussure's sense) theoretical values are used to explain absolute empirical values: *"Strange as it may seem,* the higher-level hypotheses that explain the lower-level propositions are accepted as valid purely on the basis that they do explain them, in the specific sense that they logically imply them . . . ; whereas acceptance of the lower-level propositions that need to be explained is contingent on empirical evidence . . . " (Blau 1970, p. 202; emphasis added). The strangeness, scientifically speaking, of this procedure is that the higher-level hypotheses, being abstract dictionary entries, lack a

context (logical or empirical) by which they can be explained. Thus, dictionary explanations rely upon the imposition of abstract theoretical codes on data generated by contextless measures. In other words, formalism tends to produce formal explanations and precise measurements of only one feature of organizations: formal structures which have as their distinctive feature that they are analytically context free in the same sense as the measures by which they are defined.

Therefore, it is not surprising that Blau's definition of social structure is stated in terms of its formal measurability: "[B]y social structure I refer to population distributions among social positions along various lines—positions that affect people's role relations and social interaction. This intricate definition requires explication, and *I will use the term structural parameter to clarify it*" (1974, p. 616; emphasis added). Even though he states elsewhere that "social structure is defined by the distinctions people make . . ." (1974, p. 616), the fact remains that his deductive theory is constructed entirely without reference to these decisions. In fact, he comes very close to saying that structural parameters (that is, scientist's code) create particular social behavior: "The assumption is that the differences in group affiliation and status *created by structural parameters* affect role relations and the social interaction in which these relations find expression" (1974, p. 617; emphasis added). Given other assumptions, this expression is not as elliptical as he might think (1974, p. 617, n. 5). Inasmuch as structures can only be "observed" parametrically, and since the information produced thereby is explained only by context-free hypotheses, this is equivalent to saying that, sociologically, these structures are created by their measures. This is what is meant by general theory purchased with dictionary-like abstractness. The cost is that particular social contexts are stipulated. They are not a positive source of sociological information.

Interpretative Commentaries

Bittner (1974) and Garfinkel (1967*b*)[15] are useful examples of interpretativists who have studied formal organizational structures. Bittner (1974) discusses the same problems as did Blau—the relationship among size, complexity, and administration in formal organizations:

> It is often noted that the formal organization meets exigencies arising out of the complexity and large scope of an enterprise. The rationally conceived form orders affiliations between persons and performances that are too remote for contingent arrangement, by linking them into coherent maps or schedules. This integration transcends what might result from negotiated

[15] The article cited here (Garfinkel 1967*a*), though published under Garfinkel's name, is the product of research conducted jointly with Bittner. Thus, Garfinkel (1967*a*) is taken as an empirical complement to Bittner (1974).

> agreements between continuous elements, and lends to elements that are not within the sphere of one another's manipulative influence the character of concerted action. . . . A rational principle of justice may prevail in the entire structure while governing differentially correct associations between particular performances and rewards. The varieties of demeanors that are appropriate to a particular status within the system may be perceived as variations of a more general pattern. [Bittner 1974, p. 78]

Nonethnomethodological readers of such a text often and understandably miss the point that this is merely a specific type of interpretative explanation based on data generated by a measurement technique. Two features distinguish this statement from that of Blau on the same structural topics. First, *organization* is explicitly a cognitive attribute. As such, it becomes both a topic and a resource. The concept *organization* is used to study organizations and their structures. "[T]he study of the methodical use of the concept of organization seeks to describe the mechanisms of sustained and sanctioned relevance of the rational constructions to a variety of objects, events and occasions relative to which they are not invoked" (Bittner 1974, p. 76). Second, *organization* is a cognitive attribute of both common-sense actors (Bittner 1974, pp. 74–77) and professional sociologists. As such it is highly contexted.

These two features are what lend to the above passage both its explanatory claim and its status as a context-absorbed commentary. Formal organization structures cannot be literally defined (Bittner 1974, p. 71) because they are understood, above all, as the workers' cognitive instruments for providing a "rational principle of justice" in the face of organizational complexity ("differentially correct associations between particular performances and rewards"). Thus, the text provided above works within the context of sociologists' use of the concept *organization* (Bittner 1974, pp. 69–70) in the same way as other studies (Garfinkel 1967*a*) are contexted with reference to common sense concepts of *organization* used by interviewers in a psychiatric clinic.

If, as was argued above, explanation is the theoretical encoding of already coded measurement values, interpretative commentaries are (given their own logic) explanatory. Bittner therefore uses the concept organization as "a generalized formula to which all sorts of problems can be brought for solution" (1974, p. 76). Both sociologists and organization workers are engaged in an explanatory enterprise. The only difference is that the (ethnomethodological) sociologist assumes a relatively nonprivileged responsibility to account for the explanation itself, while the actor, so to speak, simply uses the concept *organization* to measure (encode organizational reality). The crucial point is that, in this sense, all social activity is measurement (see, e.g., Cicourel 1974, p. 91).

The result of interpretative explanations of this sort is a semantics of the

very feature of formally structured organizations that formalism is forced to abandon: the particular motives of workers. Thus Garfinkel (1967*a*), in a study performed with Bittner, offers explanations for the poor quality of records kept by workers in a psychiatric clinic. These records are inadequate only when administrators, sociologists, and others with an actuarial interest bring a context-free concept of organization to them, expecting to find information useful to their (context-free) needs for comprehensive demographic data and medical histories. The records are bad only because they are not useful to formal theories of organizations. If, however, it is understood that clinical records are constructed as a therapeutic contract between the clinic and patients (Garfinkel 1967*a*, p. 199) one sees them differently. Not only are they in another sense good ("orderly and understandable" [1967*a*, p. 202]), but their value opens an entirely different structural feature of the organization to those who read them in this light. The structured relationship between the differential interests of researchers and medical practitioners becomes clear (1967*a*, p. 207).

At this point one comes to the problem that concerned Blau: the relationship among complexity, size, and administration. An organization of relatively large scale faces the organizational problem of providing actuarial records for administrative purposes (external funding, third-party payments, institutional research, etc.). This formal requirement is in conflict with a physician's need (based on ethical and legal considerations such as the threat of malpractice suits) to maintain a record of his faithful compliance with normal therapeutic procedures. Organizations thus face the cost problem. On a fixed budget it may not be possible to maintain two separate sets of records without increasing the administrative component (1967*a*, p. 192). Implicit in Garfinkel's argument is that organizations will arrive at a point where the cost of the administrative component becomes prohibitive, thus permitting bad clinical records to be kept by those with contractual rather than actuarial interests in them (Garfinkel 1967*a*, p. 200; cf. Blau 1970, pp. 214–16).

The distinctive feature of this explanatory commentary is that it measures with constant reference to workers' motives. Its deficiency is that the clarity of these measures cannot be matched by attention to general, formal features of organizations. Though Garfinkel does attend to structural features (1967*a*, p. 200), he does not do so with as much precision as Blau.

The crisis for sociologists in this incongruity between formalist and interpretativist measures is that, just as Blau can be criticized for ignoring workers' motives, so Bittner and Garfinkel can be properly criticized for slighting relatively nonmotivated features of organizations. We can be certain that the cost factor controlling the relationship between administrative and medical components is due in large part to social forces that are not available for interpretative explanation: the state of the economy, the extent

of rationalization in the social system, the availability of trained administrators, the market price of computer systems, and so forth.

Encyclopedic Texts

A common problem for both of the preceding types of explanation is motives. Formalism ignores them. Interpretativism gets caught in them. Structuralist semiotics reduces this difficulty by showing how *behavior* can be read as motivated yet not a *"product"* of motives. The interpretativist's overly strong theory of motives is an unwarranted subjectivization of *behavior.* More seriously, it cannot lead to adequate measurement because when motives are taken as the "origin" of behavior, they are mysterious and unobservable. Interpretative processes (Mead, Blumer), typifying processes (Schutz, Berger, Luckmann), rational reflexive negotiating and interpretive procedures (Garfinkel, Cicourel) are not made available to view, though their products (negotiated, concerted rational behaviors) are. Interpretativism's fear of "reducing" meaning to behavior has apparently led it to its overly strong theory of motives. A semiotic theory of explanation shows that one can measure behaviors as systematically significant expressions of motives without stipulating the latter as the exclusive cause of the former.

A semiotic measure of motives through behaviors is analogous to Saussure's linguistics which ignored the origins of speech, while using speech (*la parole*) as the behavioral products (signs) through which one reads the sense of utterances. Utterances are rendered *meaningful* not in a psychology of motives, but in a sociology of conventional linguistic structures (*la langue*). This alternative is crucial because an overly strong theory of motives as origin cannot yield a comparably strong theory of structures. When measurement is tied to motives, measurement (hence explanation) becomes metaphorical at that point at which one encounters unmotivated structures (such as the effect of the "market" on cost of administration).

The semiotic approach conceives of *"behavior"* as a *behavior function* (BF): that is, as the articulation of a motive (a significant *content*) with an *"act"* (a structurally provided *expression* unit). Motives, thus, are measured with reference to structurally provided acting *types.* Just as Saussure defined the sign of articulated speech utterances as the product of a motivated speaking individual, a BF would be the function of motivated actors. However, for Saussure, the speaking itself (*la parole*) was unmotivated insofar as the speaker has few degrees of freedom to modify and correct *la langue.* Saussure's speaking (*la parole*) is the borrowing of elements from language. Motivation therefore is assigned to the necessary but insufficient role of selection (as opposed to origination). These motivations could, in Saussurian thinking, be taken as the topics of a social psychology of speaking, but not of linguistics. Similarly, one can assign to social psychology

the task of explaining motives without assuming that this task exhausts the obligations of sociology *tout court*. Sociology explains structured behaviors.

A BF, semiotically understood, is the product of a motivated selection of unmotivated *types* of behaving. *Types* (the primary topic of sociology) are "created" not by intentions of individuals but by the conventional arrangements of the social group. Social structures provide the paradigms ("syntactical" and "lexical"—that is, norms and acts, respectively) from which actors select "normal" BF *expressions*.

I shift now from an actor's to an analyst's perspective. Social analysts would begin measurement by taking observations of BF's. It is here that I rely upon the semiotic proposals of Umberto Eco (1976).

Figure 1 presents some of the concepts discussed by Blau and Garfinkel but rewrites them in a form consistent with Eco's structural semiotics. This figure illustrates the analysis of only one "slice" of an actual BF and only several aspects of the available social structures (SS's). It is not a complete explanation. Nonetheless, one sees that BF's are observed as *expression* units (specific bad clinical records—/BCR/) which are *tokens* of a BF *type* (good organizational reasons—/GOR/). In the latter connection, it is assumed that a given bad clinical record (/BCR/) is merely the "*articulation*" of a *type* of behavior (writing bad clinical records for good organizational

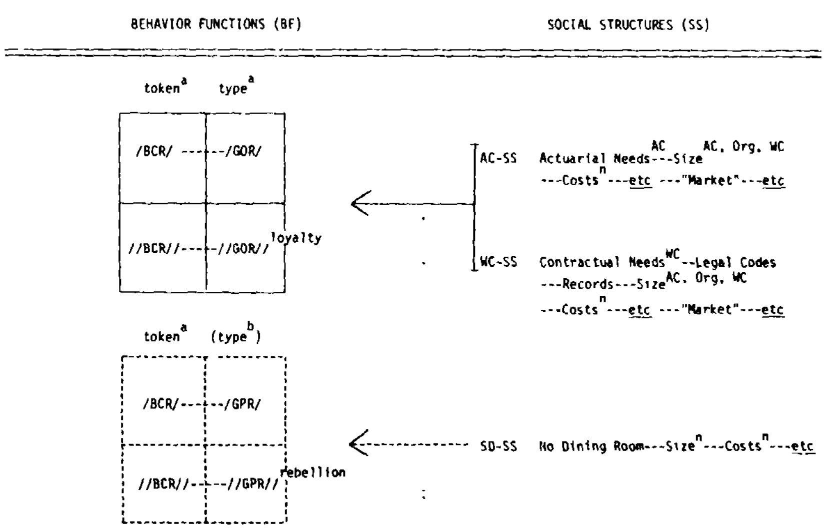


FIG. 1.—*Code:* / . . . / = expression unit; // . . . // = content unit; BCR = bad clinical records; GOR = good organizational reasons; GPR = good physicians' reasons; AC = administrative component; WC = workers' (physicians') component; *etc.* = other structural features not discussed here; other SS signs are explained in the text. Broken line around lower BF represents "possible BF not selected."

reasons—/GOR/). The /GOR/ *type* is provided by structural features of organizations (in this case, psychiatric/medical clinics). The structures, though positive in the sense that they provide *types* to BF's, are not "positivistically" positive in the sense that they are not considered available to observation. The SS's are absent to the BF's they form, just as Saussure's *la langue* is absent to the linguist observing *la parole*. The SS's are positive only in the sense that *types*—thus *token expressions*, thus BF's—can be explained only with reference to SS's. *Contents* are not the written marks on clinical records. These marks are *expressions*. *Content* describes the semantic content (*"meaning"*) of a /BCR/. It is the token contents that are measured and explained, but they are observed as *expressions* of normatively defined (that is, structured) *types*.

It is important to note that this analysis is already at a certain level of generality. I have used Bittner and Garfinkel as resources to establish the *expression type* "bad clinical records" (/BCR/). Ordinarily social research would begin at a more concrete point, prior to sociologists' knowledge of this particular BF. The "badness" of these records is already a sociologically encoded *expression type*. Normally, research would begin where Bittner and Garfinkel began one day: reading records (/BCR/s) that made little sense given the assumption of good (well-structured) formal clinical organizations.

Measurement (the explanation of the /BCR/-BF) is possible, however, only with "knowledge" of formal organizations as such: that is, their social structures (SS's). Depending upon the topic at hand and the state of sociology's structured knowledge of that topic, this knowledge need not necessarily be gained de novo. In the present case, we have good sociological reasons to "know" (from Bittner, Gouldner, Merton, Blau, and others—from the *"literature"*) that such a token BF as "bad clinical records" (/BCR/) might be "explained" (*measured*) with reference to such structural features as: exact degree of organization's differentiation (complexity), size of organization, cost of administrative component, size of administrative component, size (and cost) of worker component, and the pertinent ratio relationships linking them. I am not questioning the usefulness of these formal explanatory concepts. I am doubting that their use in either formalist or interpretativist measures can lead to an "adequate" and "complete" sociological explanation.

The major problem of an encyclopedic *measure* is evident at this point. How does one introduce a principle of selectivity? Without such a method for engendering parsimony, one would quickly be overwhelmed by the multitude of possible SS's providing *types* to even a single BF. Eco offers such a method. It is not necessary to " . . . list and to structure *all* the possible occurrences of a given item but only those which are *culturally* and *conventionally* recognized as the more statistically probable" (1976, p. 110). This would involve a preliminary explanation of the *"literature"* pertinent to the

topic at hand. In this case, *"literature"* would include sociological literatures of the organizations, formal self-descriptive literatures of the organizations themselves, workers' informal literatures (e.g., "bad" clinical records). The preliminary explanation of these literatures would yield a description of conventionally probable SS's available to provide types to the observed BF. The descriptions would be produced by means of all available and appropriate *reading* methods (descriptive statistics, analytic readings, and observational readings). This method for introducing parsimony (only partially illustrated here) is warranted by the fact that social structures are scientifically interesting only as features of sociologists' structural theories. Both interpretativists and formalists agree on this point (Garfinkel and Sacks 1970; Blau 1974).

On the other hand, the hedge against the possible abstractness of sociologists' *readings* of probable SS's is found in the fact that the readings are done empirically (*contextually*). One does not "search the literature" prior to investigation. The *readings* are over against observed BF's such as a /BCR/ *token*. The readings, however, are not hermeneutical but structural. One does not attempt to discover the "inner meaning" of the observed BF, only (at the outset) its probable structural origins.

The result of structural readings of this sort would be an encyclopedic semantic tree (such as fig. 1) displaying the SS's probably explanatory of the observed BF. Insofar as the *literatures* read are themselves encyclopedic descriptions of possible BF's, such a tree is, in effect, a second-order encyclopedia that reduces sociological and "natural" *literatures* to a relatively more formal and more general theory of the semantics of the observed BF.

It is important to note, however, that the final explanatory *"tree"* (and its accompanying text) can not select probable explanatory SS's on merely formal statistical or mathematical grounds. Instead, such selections are made comparatively, by determining the differential value of probable explanations. For example, in figure 1, the status deprivation SS (SD-SS) would describe an aspect of the structural arrangements of status prerogatives in organizations. In a clinic this could involve questions of: the "quality" (honorific appropriateness) of the medical staff's dining facilities, including presence or absence of a separate dining room; convenience of eating schedules; cost of food, heat, lighting, cooks, building space, etc.; relationship of these to overall cost factors in the organization (budget, size of administrative component, size of organization, etc.). Here the explanation of the /BCR/-BF might be that physicians write bad clinical records to rebel against "inadequate" treatment (no separate dining room because of cost factors). In other words, the SD-SS explains the //BCR// as a token of a *type:* good physicians' reasons to rebel (//GPR// rebellion).

Let us assume, however, that the "correct" explanation is along the lines

made available to us by Bittner, Garfinkel, and Blau: namely, that //BCR// is explained by size of administrative component and size of organization (assuming that patients are not members). One could only "know" of the existence of this structural branch by determining its *negative* relation to the "rebellion" SS explanatory tree (SD-SS), as well as other statistically possible explanations not illustrated here. Stated in the simplest of terms, the /BCR/-BF must be measured differentially by a comparison of, in this case, the two alternative structured *type* BF's (here loosely labeled *rebellion* and *loyalty*).

The loyalty type must then be *measured* differentially, comparing the administrative (AC-SS) and workers' (physicians' [WC-SS]) structural branches. These designations are shorthand versions of the structural details that would have to be provided (depending, of course, on the state of the *"literature"*). I am here relying (without explicit justification) on Bittner and Garfinkel's descriptions. The point I wish to make is that the //GOR// loyalty *type* BF can be established only by a description of the AC-SS and WC-SS branches. Following Blau, Bittner, and Garfinkel, one assumes that the AC-SS branch would include: (*a*) organization's actuarial needs for good clinical records, (*b*) size of administrative component, (*c*) size of organization, (*d*) relative size of worker (physician) component, (*e*) relative costs of *d* and *b*, and so forth. Likewise, the WC-SS would presumably include: (*a*) ethical and legal codes determining physicians' sense of contractual responsibility, (*b*) structured access of physicians to record-keeping procedures, (*c*) size of administrative component, (*d*) size of organization, (*e*) relative size of worker component, (*f*) relative costs of *e* and *c*, and so forth. With reference to these branches, it is clear that Bittner and Garfinkel's explanation (Garfinkel 1967*a*) requires a description of both branches. Similarly, both are necessary to incorporate motives into Blau's (1970) structural theory. Again, the explanatory value of the WC-SS is determined by its differential relationship to AC-SS.

At this point, the question is: Does this merely generate a way of rewording Bittner and Garfinkel? The test is whether or not such a semantic encyclopedia (fig. 1) would be expandable when necessary to provide access to the kind of "information" provided by Blau. Among other things, this is the question of generality. Blau studied 1,500 agencies; Bittner and Garfinkel, one. Can these two types of "information" be used in the same explanation (provided, of course, they are "properly" collected)?

Assuming that /BCR/ as a general phenomenon in clinics and agencies of all sorts is a sociologically interesting question (which it is), one could not possibly stop at the AC-SS and WC-SS explanations here described, any more than one can trust Blau's explanation without accounting for the bad records from which (in part) his data were drawn (Blau 1970, p. 204). I take just one instance: cost of administrative component relative to size of

organization. It is clear that one would indeed need to describe the *statistically probable* (Eco 1976, p. 110) structural features of administrative cost in the same manner as the illustrated SS's have been described. This would include not only cost relative to administrative and organization size (which Blau studied), but also the cost features of the WC-SS (cost of physicians' salaries, cost of recording procedures [tape recorders, computers, typing pool, etc.], "cost" of intruding on physicians' felt need to protect a contractual therapeutic relationship, etc.). In short, one can arrive rather quickly at macro questions such as "market" and "state of economy" structural conditions, structural factors that even Blau's theory omitted. The interesting point is that Blau, having chosen to ignore workers' motives, could not measure certain structural factors.

I will conclude, and briefly summarize, by referring to the way in which a semantical encyclopedia would provide for "replication." The standard sociological format for *replication* studies is a test of the extent to which interexperimental "data" support the original dictionary entries. Because of their use of *contexts*, encyclopedic explanations are insufficiently abstract to be used in this kind of comparison. However, this does not exclude the possibility of tests of inter-"experimental" reliability.

A sociologically useful semiotics of the type described here would depend on the legitimacy of understanding the absent-to-view world of social structures as a primary "empirical" *resource* available to be *"read"* as one reads an *encyclopedia*. One takes to an encyclopedia questions that are seldom answered by any single entry. To expect any one article to answer a question is to treat an encyclopedia as though it were a dictionary. Normally, encyclopedic *"answers"* are provided by a search of structured relationships appearing in the differences among aspects of "knowledge" found in different articles. The encyclopedia explains nothing itself. Explanation is given only in a *structural reading*, which leads to an explanatory *rewriting*. The rewrite (a second-order encyclopedic text) may well be stated in highly formal, even mathematical, *codes* expressing correlations, factor magnitudes, and paths.[16]

Encyclopedias are organized in a way that presupposes their usefulness in the construction of these rewrites. An encyclopedia (as opposed to a dictionary) works only because it is organized around its own internal index and cross-referencing system. Thus, *"replication"* would be based on the writing of sociological texts so that they are available for "reliable" secondary readings and rewrites. In other words, the sociological writing of an encyclopedic tree (see fig. 1) is also in part the displaying of an index of material borrowed from and, thus, cross-referenced to other *"literatures"* (sociological, everyday life, etc.). What is displayed is an explicit *code*

[16] Eco (1976, pp. 122–24) has illustrated the use of formal semantic informal processing methods for the analysis of complex *n*-dimensional semantic problems.

governing BF *types* and conventional *social structures*. This is the sense in which an encyclopedia is both context specific and general.

REFERENCES

Althusser, Louis. 1970. *For Marx*. New York: Vintage.

Attewell, Paul. 1974. "Ethnomethodology since Garfinkel." *Theory and Society* 2 (2): 179–210.

Barthes, Roland. 1970. *Elements of Semiology*. Boston: Beacon.

Bittner, Egon 1974. "The Concept of Organization." Pp. 69–81 in *Ethnomethodology*, edited by Roy Turner. Baltimore: Penguin.

Blalock, Hubert. 1968*a*. "The Measurement Problem: A Gap between the Languages of Theory and Research." Pp. 5–27 in *Methodology in Social Research*, edited by Hubert M. Blalock and Ann B. Blalock. New York: McGraw-Hill.

———. 1968*b* "Theory Building and Causal Inferences." Pp. 155–98 in *Methodology in Social Research*, edited by Hubert M. Blalock and Ann B. Blalock. New York: McGraw-Hill.

———. 1970. "The Formalization of Sociological Theory." Pp. 271–300 in *Theoretical Sociology*, edited by John C. McKinney and Edward A. Tiryakian. New York: Appleton-Century-Crofts.

———. 1973 "Thoughts on the Development of Sociology." *ASA Footnotes* 1 (2): 2.

Blau, Peter. 1970. "A Formal Theory of Organizations." *American Sociological Review* 35 (2): 201–19.

———. 1974. "Parameters of Social Structure." *American Sociological Review* 39 (5): 615–36.

Blau, Peter, and Otis Dudley Duncan. 1967. *American Occupational Structure*. New York: Wiley.

Blau, Peter, and Richard Schoenherr. 1971. *The Structure of Organizations*. New York: Basic.

Blumer, Herbert. 1969. *Symbolic Interactionism*. Englewood Cliffs, N J.: Prentice-Hall.

Chomsky, Noam. 1972. *Language and Mind*. New York. Harcourt Brace Jovanovich.

Cicourel, Aaron 1964. *Method and Measurement* New York: Free Press.

———. 1974 *Cognitive Sociology*. New York: Free Press.

Coleman, James 1964. *Introduction to Mathematical Sociology*. New York: Free Press.

Culler, Jonathan. 1976. *Saussure*. Glasgow· Fontana/Collins.

Derrida, Jacques. 1976. *Of Grammatology*. Baltimore: Johns Hopkins University Press.

Douglas, Jack. 1970 *Understanding Everyday Life*. Chicago: Aldine.

Durkheim, Émile 1938. *The Rules of Sociological Method*. Chicago: University of Chicago Press.

Eco, Umberto. 1976. *A Theory of Semiotics*. Bloomington: Indiana University Press.

Foucault, Michel 1972. *The Archaeology of Knowledge*. New York· Random House.

Garfinkel, Harold. 1967*a*. "Good Organizational Reasons for 'Bad' Clinical Records." Pp. 186–207 in Garfinkel 1967*b*.

———. 1967*b*. *Studies in Ethnomethodology* Englewood Cliffs, N.J : Prentice-Hall.

Garfinkel, Harold, and Harvey Sacks. 1970 "On Formal Structures of Practical Actions " Pp. 337–66 in *Theoretical Sociology*, edited by John C. McKinney and Edward Tiryakian. New York: Appleton-Century-Crofts.

Gibbs, Jack. 1972. *Sociological Theory Construction*. Hinsdale, Ill.· Dryden.

Gouldner, Alvin. 1954. *Patterns of Industrial Bureaucracy*. Glencoe, Ill.: Free Press.

Habermas, Juergen. 1970*a* "Summation and Response." *Continuum* 8:123–33.

——— 1970*b* "Toward a Theory of Communicative Competence." Pp. 114–48 in *Recent Sociology, No. 2*, edited by H. P. Dreitzel. New York: Macmillan.

———. 1971. *Knowledge and Human Interests*. Boston: Beacon.

———. 1973. "A Postscript to *Knowledge and Human Interests*." *Philosophy of the Social Sciences* 3: 157–89.

Haire, Mason, ed. 1975. *Modern Organization Theory*. Huntington, N.Y.: Krieger.
Homans, George. 1967. *The Nature of Social Science* New York: Harcourt, Brace & World.
Jakobson, Roman. 1974. *Main Trends in the Science of Language* New York: Harper & Row.
Kaplan, Abraham. 1964. *The Conduct of Inquiry*. Scranton, Pa : Chandler.
Lemert, Charles. In press *a* *Sociology and the Twilight of Man: Homocentrism and Discourse in Sociological Theory* Carbondale: Southern Illinois University Press.
———. In press *b*. "Ethnomethodology and Structuralism. Linguicity and the De-centering of Core Values." *Theory and Society*
Lemert, Charles, and Garth Gillan. 1977. "The New Alternative in Critical Sociology: Foucault's Discursive Analysis " *Cultural Hermeneutics* 4 (4) · 309–20.
Lepschy, Giulio C. 1970. *A Survey of Structural Linguistics*. London: Faber.
Lévi-Strauss, Claude 1966. *The Savage Mind*. Chicago University of Chicago Press.
Merton, Robert K. 1957. *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
Parsons, Talcott. 1937. *The Structure of Social Action*. Glencoe, Ill.: Free Press.
———. 1970. "Some Problems of General Theory in Sociology." Pp. 27–68 in *Theoretical Sociology*, edited by John C. McKinney and Edward Tiryakian. New York: Appleton-Century-Crofts.
Reid, Wallis. 1974. "The Saussurian Sign as a Control in Linguistic Analysis." *Semiotext(e)* 1 (2): 31–53.
Saussure, Ferdinand de. 1959. *Course in General Linguistics*. New York: Philosophical Library.
Stinchcombe, Arthur. 1968. *Constructing Social Theories*. New York: Harcourt, Brace & World.
Sudnow, David, ed. 1972. *Studies in Social Interaction*. New York: Free Press.
Wagner, Helmut. 1963. "Types of Sociological Theory." *American Sociological Review* 28 (5): 695–705.

Age and Scientific Performance[1]

Stephen Cole
State University of New York at Stony Brook

The long-standing belief that age is negatively associated with scientific productivity and creativity is shown to be based upon incorrect analysis of data. Analysis of data from a cross-section of academic scientists in six different fields indicates that age has a slight curvilinear relationship with both quality and quantity of scientific output. These results are supported by an analysis of a cohort of mathematicians who received their Ph.D.'s between 1947 and 1950. There was no decline in the quality of work produced by these mathematicians as they progressed through their careers. Both the slight increase in productivity through the thirties and the slight decrease in productivity over the age of 50 are explained by the operation of the scientific reward system. By encouraging those scientists who produce the most favorably received work and discouraging those who produce work that is not favorably received, the reward system works to reduce the number of scientists who are actively publishing. Those who continue to publish throughout their careers are a "residue" composed of the best members of their cohort. Increases in productivity through the thirties and into the forties are shown to be a result of command over the resources necessary to be highly productive. In the last part of the paper I examine the hypothesis that it should be easier for young scientists to make important discoveries in the more highly developed or codified sciences than in the less highly developed ones. The data do not support this hypothesis but, rather, suggest that scientists in all six fields are approximately equally likely to make important discoveries shortly after they receive their Ph.D.'s.

Einstein was 26 when he developed his special theory of relativity; Newton was 24 when he began his work on universal gravitation, calculus, and the theory of colors; and Darwin was 29 when he developed his theory of natural selection (Zuckerman and Merton 1973). Because of the well-known fact that many important discoveries have been made by very young scientists, it is commonly believed by scientists and laymen alike that age is negatively correlated with scientific productivity and creativity. It is as-

[1] I thank Pamela Summey for aid in data collection, Lorraine Dietrich for help with the computer, and Kenneth Bryson for statistical advice. Nancy Stern collected and analyzed the cross-sectional data on mathematics. The analysis was made easier by conversations with Ann H. Cole, Jonathan R. Cole, Lorraine Dietrich, Scott Feld, Robert K. Merton, Eugene A. Weinstein, and Harriet A. Zuckerman. This research was supported by a National Science Foundation grant (NSF-SOC 72-95 326) to the Columbia University Program in the Sociology of Science.

sumed that significant research output declines with advancing age. Perhaps the most frequently cited source of support for this assumption is the now classic work of H. C. Lehman (1953). Although Lehman's methodology has been criticized (Dennis 1956*a*, 1956*b*, 1956*c*, 1966), his conclusions are still widely accepted. It is the purpose of the first section of this paper to reexamine the relationship between age and scientific output. Is it really true that older scientists produce fewer and less significant papers than younger scientists, and, if so, why? The answers to these questions are perhaps more important today than in the past. The exponential growth of the scientific community is leveling off. This means that in the next 20 years the average age of American scientists will probably increase, as there will be relatively few new positions for young people to fill. Will the increase in age of our scientists reduce our scientific capacity? To answer this question, we must know whether the commonly held belief about the relationship between age and scientific performance is correct.

Let us begin by reviewing Lehman's work. His basic source of data is histories of science. He takes all the discoveries listed in prominent histories of science and then constructs a chart showing the number made in each five-year age period. To handle the problem of differences in life span, he uses the average number of discoveries per living scientist in his sample. His charts show that more discoveries are made by young scientists than by old ones. He does not, however, take into consideration the number of scientists alive in each age group in the population. Since science has been growing exponentially for the past few centuries, at any time the population of scientists is likely to be made up disproportionately of young people. Therefore, the conclusions based on Lehman's method are questionable.

What Lehman really has are data on only a small portion of the population of scientists—those who have made important discoveries. Instead of asking what proportion of scientists in different age groups make important discoveries, he has asked what proportion of important discoveries were made by scientists of different ages. Thus, Lehman's results are not necessarily contradictory with the conclusion that age has no causal influence on the process of discovery.

Table 1 presents a hypothetical example to illustrate the point. Suppose that at any given time we were to make a list of all the discoveries that had ever been made and were to classify them by the age of the discoverer and their importance. Most of the discoveries, important and unimportant, would have been made by young people, since science at any time has had many more young practitioners than old ones. This procedure might yield figures such as those in table 1*A*, except, of course, that the real figures would be many times larger. What Lehman has done is to examine only the first column of figures, scientists making important discoveries, and percentage in the manner illustrated in part *B*. From these percentages he would

conclude that the ability to make important discoveries declines with age. A correct percentaging of the table (shown in part *C*) would indicate that age has no influence at all on making important discoveries. Among the scientists in each age group, 10% have made such discoveries. The proper conclusion would be that the bulk of important discoveries is indeed made by young scientists because most scientists are young, not because age has a causal influence on scientific creativity.

AGE AND SCIENTIFIC OUTPUT

Two more accurate ways than that chosen by Lehman to examine the influence of age on scientific performance would be to analyze the research

TABLE 1

HYPOTHETICAL FIGURES ILLUSTRATING LEHMAN'S ERROR

A. HYPOTHETICAL DATA ON POPULATION OF SCIENTISTS

Age of Scientist	*N* Scientists Making Important Discoveries	*N* Scientists Not Making Important Discoveries	Total *N* Scientists
Under 30	500	4,500	5,000
30–39	400	3,600	4,000
40–49	200	1,800	2,000
50–59	100	900	1,000
60 or over	50	450	500
Total . . .	1,250	11,250	12,500

B. LEHMAN'S METHOD OF PERCENTAGING

Age of Scientist	% Scientists Making Important Discoveries
Under 30	40
30–39	32
40–49 .	16
50–59	8
60 or over. .	4
Total ($N = 1,250$)	100

C. PERCENTAGED CORRECTLY

Age of Scientist	% Scientists Making Important Discoveries	*N* Scientists
Under 30	10	5,000
30–39 .	10	4,000
40–49	10	2,000
50–59.	10	1,000
60 or over. . .	10	500
Total. . .		12,500

output of a sample of scientists of varying ages at one time and to examine the productivity of a cohort of scientists as they move through their careers. I have done both, but will begin by reporting the results of the cross-sectional analysis. The data consist of the research output of random samples of scientists working in Ph.D.-granting institutions. Depending on the size of the field, the samples are drawn from either all departments or every other department included in the 1969 American Council on Education (ACE) study of departmental prestige (Roose and Andersen 1970). Within each department every other member was sampled.[2] The six fields selected for study were chemistry, geology, mathematics, physics, psychology, and sociology. For each field except mathematics the sample was drawn from scientists employed in 1970 in Ph.D.-granting institutions rated by the ACE in 1969; the sample of mathematicians was drawn from those employed in ACE-rated schools in 1975.[3]

For each scientist, data on the quality and quantity of his or her current scientific output were collected. "Current output" is defined as papers published from 1965 to 1969.[4] Data on the number of papers published by each scientist were obtained from the appropriate abstracting journal. For example, the name of each chemist was looked up in *Chemical Abstracts* and a count was made of the number of papers published from 1965 to 1969.[5]

To measure the quality of publications, the 1971 volume of the *Science Citation Index* (*SCI*) was used. Counts were made only of citations to work published between 1965 and 1969, in order to arrive at a measure of the quality of current publications (Cole and Cole 1973, chap. 2).[6] By now there is a large body of literature which suggests that citations are a roughly

[2] The names of department members were obtained from either directories of graduate programs or university catalogs Emeritus professors were excluded, to limit the study to active research scientists. Scientists without Ph.D.'s, most of whom were graduate students, were also excluded in order to make the sampling procedures among departments and fields as uniform as possible.

[3] Since not all scientists who begin their academic careers at Ph.D -granting institutions will get tenure there, many scientists begin their careers in these institutions and continue them at different types of institutions. This might introduce a possible source of bias, as the older scientists in the sample had already been deemed worthy of receiving tenure, whereas the younger ones, many of whom were still assistant professors, had not yet been screened. On the other hand, it is possible that criteria for granting tenure in the major institutions have become more stringent since the end of World War II. The fact that the results from the cohort data to be reported on below are the same as those obtained from the cross-sectional data leads me to believe that any bias introduced by using a sample of scientists employed in Ph D.-granting institutions is not significant.

[4] For mathematics current output was defined as work published in the period 1970–74.

[5] All papers were included, whether they were written by the subject alone or in collaboration.

[6] Citations to the work of mathematicians published between 1970 and 1974 were looked up in the 1975 *SCI*. Past research has suggested that the failure of *SCI* to include citation data on collaborative papers of which the subject is not first author has no effect on substantive findings.

adequate indicator of the quality of work, as long as "quality" is defined as the opinion that the scientific community currently holds of the work. The number of citations to a scientist's work has been shown to be highly correlated with the way colleagues evaluate the significance of that work on questionnaires (Cole 1978), with election to the National Academy of Sciences (Stern 1978), with the likelihood of receiving a Nobel prize (Cole and Cole 1973, chap. 2), and with scientists' reports of who has exerted the most significant influence on their work (Mullins et al. 1977). As long as we are examining groups of scientists, as opposed to a specific individual scientist, there can be little doubt that those scientists who, on the average, have received a high number of citations to their work have produced work which most of their colleagues would deem more significant than those scientists who, on the average, have received relatively few or no citations.[7]

In table 2 data are presented showing the relationship between age and current productivity for the six fields, both separately and in combination. It turns out that, in general, age is curvilinearly related to productivity. Productivity rates rise, peaking either in the late thirties or the forties, and then drop off. The same curvilinear relationship was observed in five of the six fields, with the point of inflection varying slightly.[8] Similar results were

TABLE 2

AGE AND SCIENTIFIC PRODUCTIVITY, 1965–69, SIX FIELDS: MEAN NUMBER OF PUBLISHED PAPERS

	AGE						
FIELD	Under 35	35–39	40–44	45–49	50–59	60+	Total
Chemistry	11.5	11.4	16.0	16.3	13.9	10.3	13.0
	(115)	(62)	(55)	(34)	(61)	(29)	(356)
Geology	7.3	7.6	8.1	7.0	3.8	3.4	6.3
	(45)	(60)	(60)	(36)	(49)	(41)	(291)
Mathematics[a]	5.1	7.3	6.2	3.5	5.2	6.1	5.6
	(101)	(96)	(67)	(63)	(73)	(35)	(435)
Physics	4.5	5.3	6.2	5.6	4.4	3.4	5.1
	(138)	(153)	(111)	(84)	(61)	(45)	(592)
Psychology	5.6	6.4	6.4	4.9	3.3	4.4	5.4
	(151)	(101)	(92)	(94)	(79)	(27)	(544)
Sociology	1.8	3.8	5.2	6.3	5.1	4.4	4.1
	(60)	(41)	(40)	(33)	(39)	(29)	(242)
Six fields combined	6.1	6.8	7.7	6.3	5.9	4.6	6.4
	(610)	(513)	(425)	(344)	(362)	(206)	(2460)

NOTE.—N's are in parentheses.

[a] For mathematicians, 1970–74 is the period used.

[7] For a complete bibliography of studies using citations, see any one of the annual guides published by the Institute for Scientific Information to accompany the *SCI*.

[8] For mathematics we find no systematic effect of age. Mathematicians in the age group 45–49 have an unusually low publication rate. The reasons for this are not known.

obtained by Bayer and Dutton (1977), who looked at the number of self-reported publications in the two previous years by samples of scientists in seven different fields. Using regression techniques, they attempt to compute the best-fit function for each field. They conclude: "Application of linear models are especially weak in explaining variance in professional activity with career age. However, even the "best-fit" a priori model for most fields, represented by the fourth-degree polynomial, explains less than seven per cent of the variance even in the case of the field (experimental psychology) with the highest obtained R for prediction of publication output over the past two years" (1977, p. 274). Although Bayer and Dutton observe some differences in the best-fit models for the seven fields, all but chemical engineering show a curvilinear relationship between age and productivity. Since computations made on the data presented here, as well as those made by Bayer and Dutton, indicate that age does not explain much variance in productivity, I have concentrated on explaining why productivity gradually increases in the earlier years of a scientist's career and drops off at the end.

Since we are not at this point interested in field differences the data from all six fields are combined in the last row of table 2. In general, the data support Bayer and Dutton's conclusion that age explains very little variance in productivity. There is a difference of only 3.06 papers between the most productive age group (40–44) and the least productive (60+). Productivity gradually increases to about the age of 45 and then gradually decreases. In most of the fields studied the scientists over the age of 60 were not much less productive than those under 35.

Although young scientists on the average may publish fewer papers than colleagues in their forties, the papers they do publish could be more significant. Perhaps when we consider the relationship between quality of work and age we will find the expected relationship. The truly important discoveries may be disproportionately the work of young scientists. Using the number of citations received by recent work as a measure of the significance of publications a rough test of this hypothesis can be made.[9] Table 3 presents data on citations to recent work in the six fields. Again we find a basically curvilinear relationship, although the curve rises less steeply than it does for productivity. Combining data from all six fields yields the figures in the last row of the table. Basically, age has a slight curvilinear effect on the quality of published work by scientists between the ages of 30 and 50. Scientists over the age of 45 are slightly less likely to publish high-quality research than those under 45. It might be objected that the operational definition of high-quality work is not exacting enough. Perhaps the highest-quality discoveries are made disproportionately by young scientists. For the

[9] Bayer and Dutton (1977) use the total number of works cited without regard to when they were published or how many citations they received. Therefore, their data on citations are not relevant to the hypothesis we are interested in.

TABLE 3

AGE AND CITATIONS TO WORK PUBLISHED, 1965–69, SIX FIELDS: MEAN NUMBER OF CITATIONS IN 1971 *SCI*

	Age						
Field	Under 35	35–39	40–44	45–49	50–59	60+	Total
Chemistry	14.4	11.4	20.2	18.4	12.1	7.4	14.2
	(115)	(62)	(55)	(34)	(61)	(29)	(356)
Geology	5.6	6.5	7.2	5.7	2.7	1.2	5.0
	(45)	(60)	(60)	(36)	(49)	(41)	(291)
Mathematics[a]	2.7	3.8	5.8	3.4	5.6	5.1	4.2
	(101)	(96)	(67)	(63)	(73)	(35)	(435)
Physics	11.2	15.1	10.8	6.8	7.4	15.9	11.5
	(138)	(153)	(111)	(84)	(61)	(45)	(592)
Psychology	5.2	6.6	6.8	5.1	3.3	3.3	5.3
	(151)	(101)	(92)	(94)	(79)	(27)	(544)
Sociology	8	1.6	2.4	3.6	1.8	1.5	1.8
	(60)	(41)	(40)	(33)	(39)	(29)	(242)
Six fields combined	7.5	8.8	9.1	6.4	5.7	6.3	7.5
	(610)	(513)	(425)	(344)	(362)	(206)	(2460)

NOTE.—*N*'s are in parentheses.
[a] For mathematicians, 1970–74 and the 1975 *SCI* were used.

data presented in table 3, no matter how we classified "high-quality" work the result remained unchanged.

It has frequently been pointed out that longitudinal data are superior to cross-sectional data in measuring the effects of age on scientific productivity (Allison and Stewart 1974). Cross-sectional data do not allow us to disentangle the effects of age from cohort effects. Therefore, longitudinal data were collected for a cohort of mathematicians. Mathematics was chosen for this purpose because youth is believed to be more significantly correlated with creative productivity in math than in any other science (Lehman 1953). Thus, in a cohort study of mathematicians we would presumably have the greatest chance to find the anticipated effect of age not shown by the cross-sectional data. For all 497 people who earned a Ph.D. in mathematics in American universities between 1947 and 1950, data were collected on the number of papers published in each of 25 years and the number of citations to these papers.[10]

[10] The number of papers was counted in *Math Reviews;* citations were counted in the 1961, 1965, 1970, and 1975 editions of *SCI*. One problem in studying a cohort is that some people drop out as a result of death or a career change. Of the 497 people in the cohort, 52 could not be located in 1975. Some of these were probably women who married after receipt of the Ph.D. and then used their husbands' names. Of the people about whom we had information, 84% were employed in academic math departments, 7% in other academic departments, and 10% in industry. Despite the fact that some people clearly were no longer mathematicians, we did not exclude anyone from the cohort. Since people who are no longer mathematicians cannot be making scientific contributions in the later years, any bias introduced by counting all 497 people as part of the cohort goes against the null hypothesis that creative productivity is not related to age.

Productivity and citation data for this cohort are presented in table 4. As the first column indicates, the mean productivity of the members of the cohort did not change significantly during the 25-year period. In the first five-year period after the cohort had completed its dissertations, the average number of papers published was 2.4. In the last period for which we have data, 1970–74, ending 25 years after all members of the cohort had received their Ph.D.'s, the average number of papers published was 2.6. Thus, the data on a cohort of mathematicians support the conclusion based upon a cross-section analysis that productivity does not differ significantly with age. The second column of table 4 shows the mean number of citations in 1975 to papers published by members of the cohort in each of the five-year periods. The means range only from 0.84 to 1.4 for work produced in the first and fourth periods, respectively. For the fifth period the mean is 1.1. Therefore, if we used citations in 1975 as the only indicator of the quality of work, we might conclude that there have been no significant changes in the quality of work produced by this cohort throughout the 25-year period. The last column of table 4 shows the mean number of citations received by work published by members of the cohort in the different periods when we use the edition of the *SCI* which is closest in time to the publication date. For example, for the period 1950–54 the closest volume of the *SCI* would be 1961. When we use this measure we see that there has been a gradual increase in the average number of citations received to work published up until the late 1960s and then a slight decline in the early 1970s. If we were to use these citation counts as indicators of quality, we would have to conclude that on the average the members of the cohort produced more significant work in their later years than they did in the years immediately following the receipt of their Ph.D.'s.

The citations received by a given paper are a function of three different variables: the quality of the paper, the change in its relevance as time progresses, and the increase in the size of the *SCI* file. One would presume

TABLE 4

MEAN NUMBERS OF PAPERS PUBLISHED AND CITATIONS TO THEM FOR COHORT OF MATHEMATICIANS RECEIVING PH D. BETWEEN 1947 AND 1950

Date of Publication	Mean *N* Papers Published	Mean *N* Citations Listed in 1975 *SCI*	Mean *N* Citations Listed in Volume of *SCI* Closest to Time of Publication
1950–54	2.4	84	33 (1961 *SCI*)
1955–59	2 8	1 2	78 (1961 *SCI*)
1960–64	2 3	1 1	96 (1965 *SCI*)
1965–69	2.8	1.4	1.4 (1970 *SCI*)
1970–74	2.6	1 1	1 1 (1975 *SCI*)

that, as a paper gets older, on the average its relevance will decrease because the significance of the problem it addressed may decline or it may be surpassed by other work. However, since science has been growing rapidly in the 25-year period that we are studying, and the *SCI* file increased significantly in size between 1961 and 1975, the probability of a given paper being cited has been increasing as a result of this increase in the size of the file.[11] The data we have indicate that the increase in the size of the file has a far more significant effect on the number of citations received than does the possible decline in relevance of earlier work. Therefore, the least biasing measure of quality for testing the hypothesis of this paper is the use of citations in 1975 to work published throughout the 25-year period.[12]

In order to trace further what happened to members of the cohort as they progressed through their careers, a typology of productivity was developed. Each mathematician in each period was placed in one of three categories: the nonproductive, weak publishers, or strong publishers. Nonproductive mathematicians were those who published no papers at all in a given time period; weak publishers were those who published at least one paper in the five-year period but received at most one citation to their work; strong publishers were those who had published at least one paper in a five-year period and received two or more citations to the work published in that period. The data presented in table 5 tell an interesting story. In the first five-year period after receipt of the Ph.D., almost two-thirds of the members of the cohort published at least one paper. However, only 17% published papers which were cited twice or more. Throughout the time periods under consideration the proportion of the cohort which was classified as strong publishers remained almost constant, varying between 17% for the first period and 14% for the last period. However, as time progressed there was

TABLE 5

TYPOLOGY OF PRODUCTIVITY FOR COHORT OF MATHEMATICIANS RECEIVING PH.D. BETWEEN 1947 AND 1950 (%)

	Time Period				
Productivity Type	1950–54	1955–59	1960–64	1965–69	1970–74
Nonproductive	38	50	58	60	61
Weak publishers	45	35	26	25	24
Strong publishers	17	15	16	15	14
Total ..	100	100	100	100	99

[11] I am currently working out some mathematical estimates of the effects of both these variables on the citation rates over time.

[12] In fact, all the analyses of the cohort used as a measure of quality of work published throughout the period both 1975 citations and the citation index closest in time to the date of the papers' publication. All the analyses yielded virtually identical results, differing by only a percentage point or two.

a gradual dropping off of the proportion of weak publishers and an increase in the proportion of the nonproductive. Thus, by 1960–64, 15 years after the cohort had received its Ph.D.'s, 58% of them were nonproductive, 26% were weak publishers, and 16% were strong publishers. The distribution of the cohort among these types remained constant throughout the remaining periods. Of primary concern for the hypothesis under consideration is the fact that the proportion of strong publishers remained nearly constant throughout the time period. As a group these mathematicians were not likely to be more creative when they were young than they have been in their older years.

Thus far, I have shown only that the proportion of strong publishers remains constant throughout the time period. I have not addressed the question whether the mathematicians who began their careers as highly productive scholars remain highly productive or whether there is movement back and forth between the group of productive scholars and nonproductive ones. The Pearsonian correlation coefficients presented in table 6 indicate that there is a relatively high level of stability in productivity of individual members of the cohort. Since the correlation coefficients for the number of papers published by each member of the cohort in different periods vary between a low of .61 and a high of .79, we may conclude that those members of the cohort who were productive at one time were the most likely to be productive at any other time. Using the typology of nonpublishers, weak publishers, and strong publishers, I computed a five-variable table showing the location of each member of the cohort during all five time periods. I was interested in the extent to which members of the cohort would change from one category to another. Of particular interest was a switch of more than one place, that is, nonpublishers who became strong publishers or strong publishers who became nonpublishers. Only 41 mathematicians made a two-position move during all periods under consideration. Thus, less than 10% of the cohort made a significant change in their productivity patterns during the course of their career. Two hundred sixteen members of the cohort did not change their position at all, remaining nonpublishers, or weak publishers, or strong publishers consistently. Two

TABLE 6

CORRELATION MATRIX FOR NUMBER OF PAPERS PUBLISHED BY MATHEMATICIANS DURING DIFFERENT TIME PERIODS

	1950–54	1955–59	1960–64	1965–69	1970–74
1950–54		68	62	61	63
1955–59			72	74	71
1960–64				74	73
1965–69	..				79
1970–74					

hundred forty members of the cohort made changes of one position up or down; for example, a weak publisher would become a nonpublisher. Finally, a simple cross-tabulation of the mathematicians' index score at time 1, 1950–54, and time 5, 1970–74, shows that of the 71 mathematicians who were strong publishers at time 5, only one was a nonpublisher at time 1, and 34 were weak publishers at time 1. Of the 304 who were nonpublishers at time 5, only 14 were strong publishers at time 1, and 123 were weak publishers at time 1. On the basis of these data and those presented above we may conclude not only that the proportion of productive and creative scholars in the cohort remains relatively constant over time, but also that the individuals within the cohort are in general not likely to make significant changes in their productivity patterns throughout their careers. Those who start by being highly productive and creative tend to remain so as they grow older. Those who are not productive and creative early in their careers very infrequently become more productive.

For the typology used in table 5 a relatively low cutting point was employed for dividing the publishers into "weak" and "strong" ones. To be classified as a strong publisher one needed to receive a total of only two citations to all papers published during the period. It might be argued that age will have an effect on creativity only if we look at the most significant discoveries made. This hypothesis was examined by pulling out from the cohort all those members who at any time during their career had published a paper which received five or more citations in 1975. If mathematicians are indeed more creative when they are young than when they are old, we would expect the resulting sample of the 47 best papers published by members of this cohort to have been published predominantly during their earlier years. In fact, the data offer no support for this hypothesis. Of the 47 mathematicians who had published at least one paper receiving five or more citations, 22% published their most cited paper in the first five-year period, 21% in the second, 21% in the third, 23% in the fourth, and 13% in the last. It should be pointed out that some of the papers published in the last five-year period may have been published too close to 1975 to receive their full complement of citations. Thus the data based upon the most significant discoveries made by the cohort of mathematicians lead once again to the conclusion that mathematicians are no more creative in their early years than in their later years.[13] We can at least tentatively reject the commonly held belief that the creativity of scientists declines after the age of 35. In

[13] It might still be argued that even papers receiving five citations or more, although produced by only 47 people, or less than 10% of the cohort, still do not represent the truly significant breakthroughs. When the cutoff point for a significant discovery is placed at a very high level it becomes virtually impossible to study such breakthroughs statistically. In fact, when we refer to the highest level of contributions, those which introduce new paradigms, we are probably dealing with somewhat idiosyncratic contributions which are not representative of the general patterns within a scientific discipline.

fact, age seems to have very little influence on the quality and quantity of work produced by contemporary academic scientists.

SCIENTIFIC OUTPUT AND THE REWARD SYSTEM

How can we explain the curvilinear relationship of age and productivity and the slight decline in the average quality of work produced by scientists over the age of 50? I believe that these patterns are influenced by the scientific reward system. Most scientists are trained in major departments. About 65% of those studied received their doctorates from departments ranked in the top 20 in the ACE survey. Students in these major departments internalize the norm that they are expected to publish. Thus, most recent Ph.D.'s in science seek recognition by publishing. As they continue to publish, some find their work rewarded and go on to publish more. But what about those who are not rewarded? I hypothesize that those who are not rewarded are less likely to continue publishing. Thus, as a cohort of scientists advance in age the number of prolific publishers is likely to decline. Most people will not continue an activity as arduous as scientific research unless they are rewarded for it.

We can, therefore, hypothesize that since the members of an age cohort who continue to be prolific publishers after the age of 50 are the scientists who have in the past produced significant work and been rewarded, their current research will be of relatively high quality.

Data from another study can be used to demonstrate that the productivity of scientists whose early work goes unrecognized drops off. In a sampling of men and women biologists and chemists who received their Ph.D.'s in 1957 and 1958, data were collected on yearly productivity and the number of citations received in 1961. Here citations to early work are used as a measure of recognition. The data are presented in table 7. About half of the scientists started their careers by being fairly productive—publishing at least three papers in the four years after receipt of the doctorate. The work of some of them was recognized as significant (i.e., was highly cited); that of the others went unrecognized. The table clearly demonstrates that scientists whose early research efforts are cited are more likely to continue to be productive. Among the initially productive scientists, those whose early work was never cited published an average of 3.6 papers in 1966–69, whereas those whose early work received 10 or more citations published an average of 9.7 papers in 1966–69.

Data from the math cohort confirm the conclusions reached from the data presented in table 7.[14] The cohort was divided into members who published

[14] The conclusions drawn from table 7 are also supported by a study conducted by Allison and Stewart (1974). Examining the publications of a large sample of biologists, mathematicians, chemists, and physicists in the previous five years, they found that the

TABLE 7

MEAN NUMBER OF PAPERS PUBLISHED IN 1966–69, BY EARLY RECOGNITION AND EARLY PRODUCTIVITY

EARLY RECOGNITION	EARLY PRODUCTIVITY	
N Citations Received in 1961	Wrote 3 or More Papers in 4 Years after Ph D.	Wrote Fewer than 3 Papers in 4 Years after Ph D
0	3 6 (56)	1 3 (166)
1–9	5.8 (125)	1.9 (89)
10 or more	9 7 (48)	3 2 (5)

NOTE —The results were exactly the same when proportion producing four or more papers in 1966–69 was used as the measure of the dependent variable, figures in parentheses are N scientists.

one or more papers in the first five years after receipt of the Ph.D. and those who published no papers in that period. The first group was then subdivided into those who had received citations in 1961 to their early work and those who had not. The mean number of papers published by each group in the period 1960–64 was calculated. Mathematicians who had received no citations to their earlier work published an average of 2.35 papers during the period 1960–64. But those who had received one or more citations to their early work published an average of 6.55 papers during that period. Thus, once again the data show that scientists whose early work is recognized are more likely to continue to be highly productive than are others.

I was able to make one additional test of the reward system hypothesis. As a cohort of scientists ages a decreasing proportion will be highly productive, but those who remain productive should be a "residue" of the most talented members of the cohort. Therefore, we should find that the work of older scientists who are still productive should be equal or superior in quality to that of younger scientists. In order to test this hypothesis I used data from the cohort of mathematicians. In 1950–54, the first five-year period after the cohort had received its Ph.D.'s, 307 of the 497 mathematicians published at least one paper. On the average, the 307 productive scientists received 1.37 citations to their work in the 1975 *SCI*.[15] In 1970–74 only 193

dispersion of the productivity distribution was much higher for older scientists than younger ones. As age increases a smaller proportion of the scientists are responsible for a larger proportion of the total publications. Using the Gini index as a measure of inequality, they found a correlation of $r^2 = .98$ between the size of the Gini index and the number of years since the Ph.D. was obtained. This supports the conclusion that as a cohort of scientists ages there is a gradual decrease in the proportion who remain productive.

[15] If we use the 1961 *SCI* as an indicator of the quality of this work we find an average of 0.53 citations to the work of the 307 productive mathematicians.

of the 497 members of the cohort were still publishing papers. On the average, these 193 members of the cohort received 2.94 citations in 1975 to work published in 1970–75. Thus, we may conclude that the productive members of the cohort in the last period were publishing on the average higher-quality work as older scientists than were the productive members of the cohort in the early period.

The data presented above are useful in explaining the decline in average productivity and quality of work among older scientists. We must now consider why in most fields very young scientists do not publish as many papers as those in their forties. This too may be explained by the operation of the reward system. The production of scientific research is to some extent dependent upon the control of resources. Most research requires at least some funding. Certainly it is easier to conduct research leading to many publications if one has research grants, graduate students, assistants, and technical help. Because the reward system does not generally make large amounts of such resources available to scientists until they have established themselves, scientists in their late thirties and in their forties are likely to control more resources than their younger colleagues. The age-graded differential access to resources could possibly explain differences in publication rates. Although data with which to perform a direct test of this hypothesis are not available, certain data for the fields of geology and chemistry allow an indirect test. For these two fields we can examine data on the number of single-authored and multi-authored papers published by each scientist.

Scientists who work with teams of graduate students and junior collaborators are more likely to publish multi-authored papers. Therefore I hypothesized that production of single-authored papers should be less likely to increase with age than that of multi-authored papers. The data presented in table 8 offer support for this hypothesis. In geology there is a slight negative association between age and the production of single-authored papers. In chemistry both very young scientists and those over 60 publish slightly more single-authored papers than their middle-aged colleagues. In both fields, when we look at multi-authored papers we find the same curvilinear relationship as we observed in table 2. We may conclude that the effect of age on productivity is a result of the reward system "cooling out" unsuccessful scientists and making resources available to successful ones.

FIELD DIFFERENCES IN AGE PRODUCTIVITY PATTERNS

Thus far we have found that in general age is curvilinearly related to both the quantity and the quality of published papers and that this results from the operation of the scientific reward system. We have not yet examined the differences among the six scientific fields considered in this paper. Age may have a more important effect on scientific performance in some fields than in others.

TABLE 8

MEAN NUMBER OF SINGLE-AUTHORED AND MULTI-AUTHORED PAPERS PUBLISHED IN 1965–69 BY GEOLOGISTS AND CHEMISTS OF DIFFERENT AGES

	AGE					
FIELD	Under 35	35–39	40–44	45–49	50–59	60+
Geology:						
Single-authored. . . .	3.2	2.3	2.2	2.0	9	9
Multi-authored. . . .	4.1	5.3	6.1	4.9	2.9	2.6
N.	45	60	60	36	49	41
Chemistry:						
Single-authored. . .	2.3	1.5	1.7	1.6	1.2	2.8
Multi-authored .	9.3	10.0	14.3	14.7	17.8	7.6
N.	115	62	55	34	61	29

The cognitive structures of scientific fields differ. Some fields, like physics, for example, have what Kuhn (1962) calls a highly developed paradigm, while others, like some of the social sciences, are in a preparadigmatic phase or are divided into competing schools. Some sciences are highly theoretical and mathematical, while others are descriptive. Recently Zuckerman and Merton (1973) have characterized scientific fields by the extent to which knowledge is codified. "Codification refers to the consolidation of empirical knowledge into succinct and interdependent theoretical formulations" (p. 507). In highly codified fields knowledge is compacted into a relatively small number of theories which can usually be expressed in elegant mathematical language. Physics and chemistry are examples of highly codified fields, botany and zoology are less codified, and the social sciences are the least codified.

Zuckerman and Merton hypothesize that in the more codified fields like physics it should be easier for young scientists to make significant discoveries than in the less codified fields like sociology. There are two reasons for this. First is the method of gaining competence in fields of varying degrees of codification. Because in the highly codified fields knowledge is compacted, graduate students can learn quickly the current state of their field from textbooks and while still graduate students begin work on the research frontier (Kuhn 1962). In the less codified fields knowledge is not as compacted and far greater experience is needed to gain competence. "In these [the less codified fields], scientists must get command of a mass of descriptive facts and of low-level theories whose implications are not well understood" (Zuckerman and Merton 1973, p. 507). Zuckerman and Merton conclude that "codification facilitates mastery of a field by linking basic ideas in a theoretical framework and by reducing the volume of factual information that is required to do significant research. This should lead scientists in the more codified fields to qualify earlier for work at the research front . . ." (p. 510).

The second reason it is easier for young scientists to make significant dis-

coveries in the highly codified fields is variation in the social process through which discoveries are identified as significant. In the highly codified sciences there is presumably greater consensus, and it may therefore be easier to identify new discoveries as important or unimportant. In the less codified fields the identification of a discovery as significant is more dependent upon the reputation of its author. "In these less codified disciplines, the personal and social attributes of scientists are more likely to influence the visibility of their ideas and the reception accorded them. As a result, work by younger scientists who, on the average, are less widely known in the field, will have less chance of being noticed in the less codified sciences" (p. 516). For these two reasons, the ease of moving quickly to the research front and the ease in having an important contribution identified as such, younger scientists should be more likely to make important discoveries in the more codified sciences than their age peers in the less codified disciplines.

The data presented in table 3 are relevant for a rough test of the codification hypothesis. Of the six fields included, physics and mathematics are the most highly codified, followed by chemistry, geology, psychology, and sociology. If the codification hypothesis is correct, we should find a higher proportion of young scientists making important discoveries in physics, mathematics, and chemistry than in the other fields. Since the distribution of citations to scientists working in the various fields differs greatly, in order to compare scientists in different fields we must standardize the data separately for each field. All scientists with standard scores greater than zero can be classified as having made relatively high quality contributions to their fields. The proportions of scientists under the age of 35 whose work published between 1965 and 1969 received a relatively high number of citations (a positive standard score) were as follows: chemistry (29%), geology (38%), mathematics (11%), physics (25%), psychology (24%), and sociology (13%). If we examine these proportions, we find little support for the codification hypothesis. Geology is the field in which the highest proportion of young scientists has made significant discoveries and mathematics is the field in which the lowest proportion of young scientists has made significant discoveries. Furthermore, young psychologists are just about as likely to make important discoveries as young physicists and chemists.

Let us consider some possible flaws in the empirical test we have performed. It might be argued that the codification hypothesis is meant to apply only to discoveries of high import and that in the more codified fields it will be easier for younger scientists to make truly outstanding contributions. However, for the data in table 3, no matter how high we set the cutoff point for classification of important discoveries, we find no systematic differences among fields. I decided, however, to perform an additional test of the codification hypothesis and examine the ages at which the most eminent scientists in various fields make their first significant discovery.

Two sets of data were used. The first comes from questionnaires sent to random samples of university scientists in physics, chemistry, biochemistry, psychology, and sociology. The respondents were asked to name the five scientists who they felt had contributed most to their field in the past two decades. I have examined the ages at which the 10 most frequently mentioned scientists in each field made their first significant discovery. In this procedure it is important that the same criteria of what an important discovery is be used in each field. The first approximation to a more definitive test was to record the age at which each scientist published his first paper to receive five or more citations in the 1971 *SCI*. As an example, the earliest paper by C. N. Yang, the Nobel laureate in physics, receiving five or more citations in 1971 was published in 1951, when Yang was 29. There are several problems in the use of this procedure. The *SCI* contains many more chemistry, physics, and biochemistry journals in its file than it does psychology and sociology journals. It is, therefore, easier for a paper to receive a large number of citations in chemistry, for example, than in sociology.[16] To compensate for the heavier representation of hard science journals, one can use different cutoff points as to what is a significant discovery. As the cutoff point for the three natural sciences, 10 citations was used; for the two social sciences, five citations. However, as the data presented in table 9 indicate, even if the cutoff point of five citations had been used for all the fields the results would have been essentially the same.

These data offer slightly more support for the codification hypothesis. Physicists do make their first important discoveries at slightly younger ages than other scientists. However, in the relatively highly codified field of biochemistry scientists take just as long on the average to make their first important discovery as in sociology. Also, in experimental psychology, which

TABLE 9

MEAN AGE AT FIRST IMPORTANT DISCOVERY FOR FIVE FIELDS

Field	Mean Age on Publishing First 5-Citation Article	Mean Age on Publishing First 10-Citation Article	N^{a}
Physics	27	28	10
Chemistry.	30	34	10
Biochemistry	35	36	11
Experimental psychology	34		11
Clinical psychology.	34		10
Sociology	34		10

[a] In biochemistry and experimental psychology two scientists were tied for tenth place. Both were included—giving 11 scientists for these two fields.

[16] For example, in table 3 the mean number of citations to the recent work of sociologists is two, and the mean for the recent work of chemists is 14.

is presumably more codified than clinical psychology, scientists are no younger when they make their first significant contribution.

I replicated this test on another set of data: lists of the most frequently cited physicists in the 1965 volume of the *Physical Review* and the most frequently cited sociologists in the 1970 volumes of the *American Sociological Review* and *Social Forces* (Oromaner 1972).[17] These lists can be considered to be samples of the scientists currently producing the most useful work in each of the fields. At what age did each scientist make his or her first significant contribution? Using the cutoff point of five citations I found that the mean age at which the physicists made their first important contribution was 31, as compared with a mean age of 34 for the sociologists (see table 10).[18]

It is possible, however, that the results might be influenced by the difference in age at which physicists and sociologists earn their doctorates. The mean age at which the physicists earned their doctorates is 26, as compared with 28 for the sociologists. Although it might be claimed that this merely reflects the shorter time required to learn the more codified field, it also might be a result of the different requirements for the Ph.D. In physics doctoral dissertations generally concern shorter, more discrete, problems than dissertations in sociology, which are usually more equivalent to books than articles. At any rate, scientists in both fields generally have little opportunity to publish prior to receipt of the Ph.D. Since physicists are younger at receipt of the doctorate, to make the two fields more comparable we should examine the professional age (years since earning the doctorate) at which scientists in the two fields made their first important contribu-

TABLE 10

MEAN AGE AT MAKING FIRST SIGNIFICANT DISCOVERY FOR MOST-CITED PHYSICISTS AND SOCIOLOGISTS

	Physics	Sociology
Mean age at publication of 5-citation discovery	31	34
N	50	31
Mean professional age	5.0	6.0
N	50	31
Mean professional age: scientists under 60 in 1971	4.9	4.8
N . . .	48	24
Mean professional age: scientists under 50 in 1971	3.9	3.4
N . .	42	17

[17] The data for the *Physical Review* were provided by Cullen Inman of the American Institute of Physics.

[18] Using the same cutoff point for each field should compensate for any possibility that work in physics might become obsolete more quickly than work in sociology.

tion.[19] It took the physicists on the average 5.0 years from receipt of the doctorate and the sociologists 6.0 years to make their first significant contribution. Since citations made in 1971 may not be a very reliable indicator of the quality of very old work, we can omit from the sample all scientists 60 years old or older in 1971. The mean professional age at which the physicists under 60 made their first important contribution is 4.9, as compared with 4.8 for the sociologists. When scientists 50 or older in 1971 are excluded we get a mean professional age of 3.9 for physicists and 3.4 for sociologists. No matter which comparisons we make, we must at least tentatively conclude that there is no significant difference in the age at which scientists in a highly codified field like physics and a relatively uncodified one like sociology make their first significant contribution.

CONCLUSION

The studies reported in this paper suggest that the relationship between age and scientific performance is influenced by the operation of the reward system. In other research I have found that the different scientific fields have very similar reward systems. The reward systems of highly codified fields like physics and chemistry do not differ significantly from those of less codified fields like psychology and sociology. The homogeneity of reward systems is influenced by the fact that the central locus for the various scientific fields is the university. Both physicists and sociologists must be evaluated for promotion and tenure within the same time period (usually five to seven years). Thus, all university-based scientists must make a significant contribution to their discipline relatively quickly if they are to maintain their positions.

A comparison of the ages at which scientists in different fields make their first important discovery leads to the suggestion that the Zuckerman and Merton (1973) codification hypothesis must be modified to take into account the type of discovery made. In some fields, for example, it may be easier for young scientists to make significant theoretical contributions. Also, the age at which synthesizing contributions are made may vary with the level of codification. In testing the modified codification hypothesis we should be careful to look at the distribution of various types of discoveries as well as the ages at which they are made. If, for example, the proportion of theoretical discoveries made in physics is greater than that made in sociology, it would not be surprising to find more young physicists making theoretical discoveries than young sociologists. A careful analysis of the age distribution of different types of discoveries should be the next step in the

[19] If a contribution was made prior to receipt of the doctorate, it was counted as a negative number in computing the averages. Only five of the physicists and one of the sociologists published a five-citation discovery prior to receiving the doctorate.

study of the influence of cognitive content on the relationship between age and scientific performance.

All the data reported in this paper lead to the conclusion that, at least in the six fields we have studied, age has only a minor influence on scientific performance. Those scientists who begin their careers by publishing and are recognized for making important contributions generally continue to publish high-quality work. It is unlikely that an increase in the mean age of our scientists will in and of itself bring about a meaningful decline in our scientific capacity.

REFERENCES

Allison, Paul D , and John A. Stewart 1974. "Productivity Differences among Scientists· Evidence for Accumulative Advantage." *American Sociological Review* 39 (August): 596–606.

Bayer, Alan E., and Jeffrey E. Dutton. 1977. "Career Age and Research—Professional Activities of Academic Scientists: Tests of Alternative Non-linear Models and Some Implications for Higher Education Faculty Policies " *Journal of Higher Education* 48, no. 3 (May/June): 259–82.

Cole, Jonathan R., and Stephen Cole. 1973. *Social Stratification in Science*. Chicago: University of Chicago Press.

Cole, Stephen. 1978. "Scientific Reward Systems: A Comparative Analysis " Pp 167–90 in *Research in Sociology of Knowledge, Sciences and Art*, edited by R. A. Jones. Greenwich, Conn.: JAI Press.

Dennis, Wayne 1956*a*. "Age and Achievement: A Critique " *Journal of Gerontology* 11, no. 2 (July): 331–37.

———. 1956*b*. "Age and Productivity among Scientists." *Science* 123 (April 27): 724–25.

——— 1956*c*. "The Age Decrement in Outstanding Scientific Contributions " *American Psychologist* 13 (August): 457–60.

———. 1966. "Creative Productivity between the Ages of 20 and 80 " *Journal of Gerontology* 21, no. 1 (January): 1–8.

Kuhn, Thomas S. 1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lehman, Harvey C 1953. *Age and Achievement*. Princeton, N.J.: Princeton University Press.

Mullins, Nicholas C., Lowell L. Hargens, Pamela K. Hecht, and Edward L. Kick. 1977. "The Group Structure of Cocitation Clusters." *American Sociological Review* 42 (August): 552–62.

Oromaner, Mark. 1972. "The Structure of Influence in Contemporary Academic Sociology." *American Sociologist* 7 (May): 11–13.

Roose, Kenneth D., and Charles J. Andersen. 1970. *A Rating of Graduate Programs* Washington, D C : American Council on Education.

Stern, Nancy. 1978. "Age and Achievement in Mathematics." *Social Studies of Science* 8 (February): 127–40.

Zuckerman, Harriet A., and Robert K. Merton. 1973. "Age, Aging and Age Structure in Science " Pp. 497–559 in Robert K. Merton, *The Sociology of Science: Theoretical and Empirical Investigations*, edited by Norman Storer. Chicago: University of Chicago Press.

Education and Occupational Mobility: A Comparison of Polish and American Men

John W. Meyer and Nancy Brandon Tuma
Stanford University

Krzysztof Zagórski
Polish Academy of Sciences

We use recent survey data to compare intergenerational mobility of Polish and American men. Father-to-son occupational mobility is considerably lower in Poland than in the United States, mainly because Poland has a large, traditional agricultural sector with high occupational inheritance. Furthermore, education, which is substantially affected by father's occupation in both countries, has a much larger effect on occupational attainment in Poland. This striking difference is not so surprising when one considers that the Polish educational system has stronger allocative powers, is more highly selective, and is differentiated laterally (into types) as well as vertically (into levels). When farmers' sons are excluded, the overall rate of mobility in Poland is similar to that in the United States. However, within the urban sector the direct effect of father's occupation on son's occupation almost disappears, while educational effects remain strong. So Polish urban mobility is almost completely filtered by the educational system.

In this paper we use data from recent large-scale surveys to compare father-to-son occupational mobility in Poland and the United States. Our main interest lies in showing the different role that education plays in the two countries. Poland has a larger and more traditional agricultural sector than the United States, a more selective and differentiated educational system, and a political system that more consistently promotes equality of opportunity and the use of education to allocate occupations. We compare the two countries to see what differences these factors make.

The literature provides some basic issues and expectations that motivate our comparative analyses:

1. We expect to find similar mobility rates in the two countries because

of the research tradition showing small differences among relatively developed countries.

2. Education may affect occupational status in Poland more strongly than in the United States. Poland's selective educational system (one of sponsored mobility; see Turner [1960]) and centrally planned economy, which emphasizes occupational selection in terms of education, work toward this end.

3. Father's occupation may have similar effects on son's educational attainment in the two countries. The Polish system emphasizes equality of opportunity more, but its early educational selection and traditional status differences (especially between urban and rural groups) work the other way.

4. The structure of the Polish educational system is more differentiated than the American one. Most secondary students enter basic vocational schools—two- or three-year programs that do not furnish direct access to higher education. Only a minority attend the four-year university preparatory or technical programs that have such access. So we expect the inclusion of a variable describing the type of secondary school attended to (1) increase the prediction of occupational attainment in Poland and (2) lower the direct effect of father's occupation in that country. Polish higher education is also differentiated into types, but our crude occupational categorization does not permit us to detect its effect on attainment. Differential programs within secondary schools, colleges, and universities may well have further effects in both countries, but we lack data to study these.

DATA

We focus on the mobility of men. For the United States we employ the 1972–76 NORC General Social Surveys on the noninstitutionalized population age 18 and over. These data have been shown to parallel other contemporary studies of American mobility (see, e.g., McClendon 1977). We include all men with complete educational and occupational data.

The Polish data come from a 0.5% random sample of the economically active population conducted by the Central Statistical Office in 1972 (Zagórski 1976*a*, 1976*b*). We exclude the few (4% of all men) who were (or had fathers who were) economically inactive, of unknown occupation, employed by entrepreneurs in their own home as "outworkers," or employed as commission agents.

The available Polish data do not permit us to control for other relevant variables, such as son's age. However, we do know that most sons (72%) completed their schooling and took their first jobs after the 1945 reorganization of Polish society.

MEASURES

Occupation

Sons and fathers are classified as having nonmanual (scored 3), manual (2), or farm (1) occupations. Ordering of occupational categories (customary in American research) simplifies our analyses but does some violence to current Polish sociological thinking, which sees these groups as differentiated rather than stratified (Zagórski 1976*a*, 1976*b*). We code sons' current occupations. Fathers' occupations are coded as of the son's sixteenth birthday for the United States and as of the son's first permanent job for Poland.

Certain limited exceptions occur in coding of occupations in Poland. First, employees of state farms and members of agricultural cooperatives (together only about 10% of the male agricultural work force) are coded as manual workers. Second, a category of independent workers, who are typically self-employed artisans or small businessmen, are coded as having nonmanual occupations, which reasonably reflects their current social position in Poland. This category includes about 1.6% of the current Polish work force but was larger in the past.

Educational Level

We constructed a scale that roughly approximates the standard classification of years of schooling. The cross-national equivalents are defined in table 1B.

Educational Type

A dummy variable was constructed to compare Polish men whose highest school was basic vocational school (scored 0) with all others (scored 1). This variable is only used in analyses that include educational level, so in effect it compares men who attended (but did not complete) university preparatory or technical schools with those who attended basic vocational school, as these two groups have about the same educational level (in terms of years of schooling).

RESULTS

Overall Mobility

Table 1A reports the basic mobility data for the two countries. We find: (1) Overall, there is more occupational inheritance and less mobility in Poland than in the United States ($r = .44$ for Poland, .38 for the United

TABLE 1

OCCUPATIONAL MOBILITY OF ADULT MALES. POLAND AND THE UNITED STATES

A. SON'S OCCUPATION BY FATHER'S OCCUPATION

FATHER'S OCCUPATION	SON'S OCCUPATION POLAND				FATHER'S OCCUPATION	SON'S OCCUPATION: UNITED STATES			
	Non-manual	Manual	Farm	Total N(= 100%)		Non-manual	Manual	Farm	Total N(= 100%)
Nonmanual	55.2	41.2	3 6	4,273	Nonmanual	66 5	32.9	5	762
Manual	22 2	72.3	5 5	13,122	Manual	35 4	63 8	8	1,483
Farm.	13 3	41 3	45 4	19,110	Farm...	24 4	55 7	19 9	734
Total..	21 4	52 4	26 2	36,505	Total ..	40.7	53 9	5.4	2,979
		r= 441					r= 376		

TABLE 1 (*Continued*)

B. Son's Occupation by Son's Educational Level and Father's Occupation

Father's Occupation and Son's Education	Son's Occupation, Poland: Non-manual	Manual	Farm	Total N(= 100%)
Nonmanual:				
University complete	99.5	.5	0	790
University incomplete	92.3	7.3	.4	248
Secondary complete	81.6	17.3	1.1	1,195
Secondary incomplete	20.1	78.1	1.9	1,186
Primary complete	15.8	74.0	10.2	722
Primary incomplete	11.4	55.3	33.3	132
Manual:				
University complete	98.5	1.1	.4	472
University incomplete	94.6	5.4	0	222
Secondary complete	75.4	23.6	.9	1,841
Secondary incomplete	11.1	87.4	1.5	4,564
Primary complete	6.5	86.6	6.9	5,074
Primary incomplete	1.2	68.9	29.9	949
Farm:				
University complete	98.4	0	1.6	432
University incomplete	95.6	1.5	3.0	203
Secondary complete	77.9	14.9	7.2	1,269
Secondary incomplete	14.0	65.8	20.2	2,782
Primary complete	5.3	49.8	44.9	9,618
Primary incomplete	.6	22.6	76.9	4,806

Father's Occupation and Son's Education	Son's Occupation, United States: Non-manual	Manual	Farm	Total N(= 100%)
Nonmanual:				
University complete	90.8	8.8	.4	262
University incomplete	67.1	32.9	0	216
Secondary complete	51.2	48.3	.6	172
Secondary incomplete	33.8	66.2	0	74
Primary complete	33.3	61.9	4.8	21
Primary incomplete	23.5	70.6	5.9	17
Manual:				
University complete	84.8	14.7	.5	204
University incomplete	53.7	46.3	0	272
Secondary complete	25.9	73.2	.8	474
Secondary incomplete	18.3	80.7	1.0	301
Primary complete	14.7	85.3	0	109
Primary incomplete	9.8	87.0	3.3	123
Farm:				
University complete	80.0	10.0	10.0	70
University incomplete	46.0	38.1	15.9	63
Secondary complete	25.4	53.8	20.7	169
Secondary incomplete	12.2	74.0	13.7	131
Primary complete	14.0	59.5	26.4	121
Primary incomplete	10.0	65.6	24.4	180

NOTE.—Four or more years of college in the United States = university completion in Poland = a score of 6. Three or fewer years of college in the United States = university not completed, or college completed or not completed in Poland = 5. Twelfth grade completed in the United States = Polish university preparatory or technical secondary school completed = 4. Nine through 11 grades completed in the United States = Polish university preparatory or technical secondary school not completed, or basic vocational school completed or not completed = 3. Eighth grade completed in the United States = Polish primary school completed = 2 Less than eighth grade completed in the United States = Polish primary school not completed = 1.

States). (2) This is so mainly because of the much larger agricultural sector in Poland. When sons of farmers are excluded, occupational inheritance declines (more sharply in Poland), and the two countries seem quite similar ($r = .28$ for Poland, .29 for the United States).

One view is that there is lower mobility in Poland. But an equally appropriate interpretation is that there was greater initial rural-urban inequality in Poland. In their social and economic situation Polish farmers more closely resemble a traditional peasantry than capitalistic American farmers. For instance, in 1970 only 3.5% of individual Polish farmers owned more than 15 hectares (37 acres) of land.

Educational Level

Table 1B reports father-to-son mobility when sons are classified by educational level. Clearly the effects of educational level are larger in Poland than in the United States. The difference in attainment created by completion of university preparatory or technical school is particularly striking.

The relationships in the data are summarized in the regression results reported in table 2. Equations (1–4) show:

1. The effect of son's educational level on occupational attainment is much higher in Poland (eqq. [3] and [4]). Two factors seem to be involved. First, the Polish nonagricultural economy is almost completely socialized and centralized. Rules of recruitment to occupational positions, which emphasize education, are strictly observed. Thus the standardized slope for education is nearly 50% higher for the Polish data than for the American. Second, the Polish educational system is much more selective, so that each of the more advanced levels of education consists of a more thoroughly screened group (with correspondingly greater occupational advantages) than in the United States (Boudon 1974). As a result, the unstandardized slope for education is 100% higher in Poland than in the United States. The higher standardized slope of education in Poland indicates the greater allocative powers of the Polish educational system; the even greater relative difference in the unstandardized slopes shows the effect of the greater selectivity of Polish education.

2. The direct effect of father's occupation on son's occupation—which we expected to be smaller in Poland—is about the same in the two countries (eqq. [3] and [4]). However, we show below that the expected differences do appear in the urban sector.

3. The effects of father's occupation on son's educational level are similar in the two countries (eqq. [1] and [2]). The unstandardized slope is slightly higher in the United States because its educational system is more expanded than the Polish one.

TABLE 2

REGRESSION ANALYSES OF MALE OCCUPATIONAL MOBILITY POLAND AND THE UNITED STATES

Equation	Dependent Variable	Father's Occupational Level	Son's Educational Level	Type of Secondary School Son Attended	R^2	Cases
(1)	Poland: son's ed. level.	73 (41)			.171	36,505
(2)	U.S.: son's ed. level	86 (40)			157	2,979
(3)	Poland: son's occ. level	20 (20)	33 (59)		.481	36,505
(4)	U.S.: son's occ. level	18 (22)	.15 (41)		279	2,979
(5)	Poland: son's occ. level	21 (21)	34 (60)	17 (10)	490	36,505
		Sons of Nonfarm Fathers Only				
(6)	Poland: son's ed. level	97 (34)			114	17,395
(7)	U.S.: son's ed. level	.89 (30)			090	2,245
(8)	Poland: son's occ. level	.08 (06)	.28 (64)		437	17,395
(9)	U.S.. son's occ. level.	17 (16)	16 (44)		.261	2,245
(10)	Poland: son's occ. level.	05 (.04)	.28 (65)	.25 (21)	479	17,395

NOTE.—We report unstandardized regression coefficients with standardized regression coefficients in parentheses Due to the large sample size, all coefficients are significant at the .001 level. The same analyses using Somers's *d* coefficients parallel those reported here. We do not report the effect of father's occupation on type of school attended. Given the coding of type of school, this relationship is nonmonotonic. Sons of manual workers are much more likely to attend basic vocational schools than sons of either farmers or nonmanual workers, who differ greatly in their educational level. Type of school, because of the way we coded it, is meaningful only when included as an independent variable along with educational level

4. Because educational level affects occupational attainment more in Poland than in the United States, more variation in occupational attainment is accounted for in the Polish analysis than in the American one (48.1% versus 27.9%).

Thus far, then, we conclude that the lower overall mobility in the Polish data is caused by two factors. First, a large part of the Polish labor force is concentrated in a traditional agricultural sector with high occupational inheritance. Second, the greater effect of education on occupational attainment in Poland (given the effect of father's occupation on son's education) indirectly produces greater occupational inheritance.

Type of School Attended

Equation (5) of table 2 shows the results of including a dummy variable for attendance at a basic vocational school (scored 0) in comparison to all others (scored 1). We see that:

1. Type of school attended substantially affects occupational attainment in Poland. The raw data (not presented) show that the great majority of students attending basic vocational schools end up in manual occupations, while those entering university preparatory or technical schools (whether to completion or not) end up in nonmanual occupations. This substantial effect—obvious to Polish sociologists, who know that it describes the intended roles of the school types—may give pause to those who imagine that the small school effects in the United States are paralleled everywhere. Obviously, when schools are given different allocative powers in a society, their effects can be different (see, e.g., Kerckhoff [1975] for similar findings on English education). Note, however, that a complete analysis of school effects requires more controls (e.g., ability, academic performance, initial aspirations) than are possible with available data.

2. The overall increase in predictive power of the analysis in equation (5) over that in equation (3) is slight ($R^2 = .490$ versus .481).

3. The inclusion of school type, contrary to our expectation, does not reduce the direct effect of father's occupation on son's occupation, which remains about as large as in the American case. However, this turns out to reflect the larger agricultural sector in Polish society.

When we eliminate sons of farmers from the analyses (see eqq. [6] through [10] of table 2), a different picture emerges. The results are striking. The effects discussed earlier remain, with an important exception. The direct effect of father's occupation on son's occupation in the simpler Polish analysis (eq. [8]) is very small—considerably smaller than in the comparable American analysis (eq. [9]) or in the Polish analysis including

sons of farmers (eq. [3]). But the effect of education on the attainment of a nonmanual occupation remains substantially larger in Poland (eqq. [9] and [10]).

These analyses show that the inclusion of type of school scarcely alters the effect of educational level (cf. eqq. [8] and [10]). So the inclusion of type of school substantially increases the overall prediction of occupational attainment: explained variation rises from 43.7% to 47.9%.

CONCLUSIONS

Occupational mobility is lower in Poland than in the United States, mainly because agricultural occupations in Poland are numerous and commonly inherited. Father's occupation has similar effects on educational attainment in the two countries, but the effect of education on occupational attainment is substantially larger in Poland. We find this plausible because Polish education is organized into a system of types as well as levels, utilizes early selection, and is intended to function as a basis for occupational allocation. Within the Polish urban sector, the differentiated, selective educational system almost completely mediates the effects of father's occupation on son's attainment, while within the United States the more fluid educational system allows substantial direct effects to remain.

Two changes in the structure of Polish society may alter Polish mobility. Urbanization should lower direct and overall occupational inheritance (cf. eqq. [3] and [8]). And a planned shift to less differentiated secondary schooling may increase direct status inheritance (cf. eqq. [8] and [10]). Still, the latter change will probably lower overall inheritance by lowering the effects of education on attainment.

The value of our analysis is limited by the available data. We have used crude classifications of occupations and have been unable to control age and other relevant variables. So our results should be treated primarily as hypotheses to be verified in later, more refined analyses.

REFERENCES

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality.* New York: Wiley.

Kerckhoff, A. C. 1975. "Patterns of Educational Attainment in Great Britain." *American Journal of Sociology* 80 (May): 1428–37.

McClendon, McKee J. 1977. "Structural and Exchange Components of Vertical Mobility." *American Sociological Review* 42 (February): 56–74.

Turner, Ralph H. 1960. "Sponsored and Contest Mobility and the School System." *American Sociological Review* 25 (December): 855–67.

Zagórski, Krzysztof. 1976*a* *Zmiany Struktury i Ruchliwość Spoleczno-Zawodowa w Polsce.* Warsaw: Glowny Urzad Statystyczny.

———. 1976*b*. "Changes of Socio-Occupational Mobility in Poland." *Polish Sociological Bulletin* 2 (34): 17–30.

Commentary and Debate

The comments printed in this section are limited to brief critiques of articles and reviews published in the *AJS*. These comments are expected to address specific errors or flaws. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Authors of articles and reviews are invited to reply to comments, subject to the same length limits. The *AJS* does not publish commenters' rebuttals to authors' replies. Longer or less narrowly focused comments on articles published in the *AJS* should be submitted themselves as articles.

COMMENT ON CAIN'S REVIEW OF *RESEARCH ON HUMAN SUBJECTS*

We are pleased that *AJS* has finally, five years after its publication, reviewed our book, *Research on Human Subjects: Problems of Social Control in Medical Experimentation* (*AJS* 84 [November 1978]: 753–55), which is now going into its first paperback edition. However, we are surprised and disappointed that the reviewer, Leonard Cain, could have assumed, as he does, that we would have done this study—or that our medical researcher subjects would have allowed us to do it—without adequate voluntary consent.

In the book, we indicate that our National Survey questionnaire was sent to the individual in each institution with whom the Public Health Service (PHS) had negotiated an assurance of compliance with PHS guidelines for the protection of the rights and welfare of human research subjects (p. 14). We also note, though we do not elaborate, that the Two-Institution Interview Study was explained to each prospective researcher-respondent in a cover letter accompanying the questionnaires (p. 25). Cain could easily have found out from us that in our cover letters we explicitly told the respondents the following: "As you are probably well aware, the whole area of human experimentation has come under vastly increased discussion in the past couple of years. There has been, however, no systematic empirical research attempted which aims at discovering facts that would better inform

the whole discussion. The enclosed questionnaire is the first part of some extensive research designed to remedy that situation." In the letter requesting personal interviews, we spoke of "the difficult, complex, but very important problem of the use of human subjects in biomedical research." In the context of the cover letter, the questions that take up a majority of the pages in each of the two schedules (appendices 1 and 2 in the book) plainly relate to the institutional review committee concerned with the protection of research subjects' rights and welfare and to the ethical issues of risk, benefit, or consent.

In both of our studies, we were given much positive evidence by the respondents that the goal of humane therapy and the problems of establishing proper procedures in human experiments were of considerable importance. For example, in the course of their interviews, many respondents were eager to know what we had learned so far, how "most researchers had answered," how typical their own responses were. We were not only requesting cooperation therefore; we were offering something in return—information about the solution of ethical dilemmas faced by our respondents. Indeed, we were so concerned about the informed-consent issues that one of us has written a special paper on the problems of access to a powerful profession (Barber 1973).

Cain argues that "there is a likelihood that many of the respondents placed the hypothetical cases in scientific, not ethical, contexts." This is very doubtful in light of the above. Moreover, it overlooks the fact that in each hypothetical research proposal there were presented not only scientific information but also one or more ethical considerations—risk, benefit, or consent, or a combination of these; and for each proposal the respondent was asked, "in the hypothetical case that you constituted an institutional review 'committee of one,' what decision would you make [regarding approval] after reviewing the proposal" (see pp. 214–25, 244–46). Our construction of the questions in this way was intended to, and most likely did, forestall the tendency of respondents, which Cain suggests, to leave ethical judgments "to the senior researcher and higher monitoring boards."

In addition, Cain charges that "ethical standards are presented as absolutes by the authors." In stating this, he ignores the discussion in which we emphasize that "the studies or respondents in our two samples were defined as being *more or less* 'favorable,' as being *more or less* 'strict' or 'permissive,' *in comparison to one another, not simply by reference to some absolute ethical standards*" (p. 52; see also pp. 44, 72–73, n. 22).

We also regret what Cain's remark, "If it had been made clear to respondents that they were being subjected to investigation of their personal ethics," implies. We not only have assured and maintained absolute confidentiality to both the institutions and the individuals who cooperated in our study (p. 25), but we have made it clear that we were not investigating

the ethics of particular individuals. It should have been clear to Cain from our book that one of the main features of our research was interest, not in individual delinquencies, but in patterned rates of delinquency and the sources of these rates in such social system origins as the science–humane therapy dilemma, inadequate medical education, inadequate peer-review control, and inadequate local institutional review control.

We sympathize with Cain's point that social researchers have often been as inadequate as biomedical researchers in observing ethical responsibilities to subjects. Indeed, Barber made just this criticism in "Liberalism Stops at the Laboratory Door" (Barber 1975). There are still plenty of suitable targets for Cain. He just made a mistake this time.

BERNARD BARBER
Barnard College

JOHN J. LALLY
Lehman College, City University of New York

JULIA LOUGHLIN MAKARUSHKA
Syracuse University

DANIEL F. SULLIVAN
Carleton College

REFERENCES

Barber, Bernard 1973. "Research on Research on Human Subjects. Problems of Access to a Powerful Profession." *Social Problems* 21 (Summer)· 103–12.

———. 1975. "Liberalism Stops at the Laboratory Door" Paper presented at the annual meeting of the American Sociological Association, San Francisco, August 25–29

REPLY TO BARBER ET AL.

The authors of *Research on Human Subjects* suggest that I "could easily have found out from [them]" that they told their respondents that "the whole area of human experimentation has come under vastly increased discussion in the past couple of years." They also insist that they obtained "adequate voluntary consent" from their respondents. After studying closely the information that the authors obligingly would have shared with me and after reading the entire book again, I still find no statement which suggests that respondents were informed that the study would lead to published conclusions about the ethics of those interviewed.

The authors continue to share information not published in their book, to wit: "[M]any respondents were eager to know what we had learned so far." If I had been responsible enough to inquire about this point, I could have learned that the authors "were offering something in return—informa-

tion about the solution of ethical dilemmas." But the book, as reported in my review, states that a significant minority of the respondents revealed an "unawareness of the importance of, or concern with, consent."

I acknowledge that the respondents were asked to think of themselves as a "committee of one" to consider approval of the several hypothetical proposals presented. The point to be made is that, since the book itself reported that ethical issues were frequently handled by senior researchers, the junior researchers, even if asked to assume the unrealistic role of a committee of one, may well have been content to concentrate only on the methodological merit of the proposal.

In the hypothetical cases, the respondents were not asked to be themselves and were not asked, to my knowledge, to concentrate on the ethical issues enfolded in the case descriptions. Let me suggest that unrealistic roles applied to hypothetical cases may not generate very meaningful data for sociology, especially when case studies could have been included. (There is, however, a personal embarrassment on this issue. In imagining myself to be a junior researcher reacting to the hypothetical cases, I allowed myself to defer to senior researchers, i.e., to validate the Eichmann defense. I should not have accepted a dichotomy between science and ethics.)

If there is ambiguity in my reference to "personal ethics," let me clarify. There was surely no accusation that the authors planned to nail each respondent's "ethics score" to the office door. Rather, I meant to suggest the likelihood that, if respondents had been asked directly about issues of informed consent, risk, etc., and had been told that certain answers were to be labeled unethical, their responses would have been less "unethical." The issue is informed consent, not confidentiality.

Again, in interpreting the issue of "ethical absolutes," my remarks have been misconstrued. The authors state that "no [absolute ethical] standards exist." They also accept a consensus sufficient to speak of a "moral community" among biomedical researchers. The point I sought to make was that, although the data raised serious questions about a moral community, the authors persisted in labeling a high percentage of the respondents as unethical. A strong case may, indeed must, be made for ethical absolutes, especially in biomedical research. But confusion is produced when absolutist labels are imposed although no absolutes are acknowledged, and although respondents were not asked their own views on the ethical dimensions of their responses.

The authors' response to my review begins with a declaration of disappointment that I could have assumed that they would have conducted their study "without adequate voluntary consent." My naiveté, before reading their book, would probably have led me to accept the proposition that most biomedical researchers obtain adequate voluntary consent before proceeding with research. This book dispelled my naiveté. Do Barber et al. want me to

resume it when sociologists conduct research? They conclude by suggesting I made a mistake. I remain unconvinced.

LEONARD D CAIN

Portland State University

COMMENT ON DUSTER'S REVIEW OF *AFFIRMATIVE DISCRIMINATION*

I would like to make two points in response to Troy Duster's review of my *Affirmative Discrimination: Ethnic Inequality and Public Policy* (*AJS* 84 [July 1978]: 252–56), one on empirical issues, one on policy issues.

On the empirical point, I briefly review census data and econometric studies which suggested to me that we no longer faced, in the early 1970s, an issue of massive discrimination against blacks. The "quotation" that I endorse is no mere quotation: It is the key conclusion of an econometric analysis by Robert E. Hall and Richard A. Kasten (1973). More important for this issue is an earlier analysis by Richard B. Freeman (1973), from which I quote more fully, and which came to the same conclusion: It was not possible to demonstrate discrimination against young blacks in the labor force in the early 1970s. Freeman's analysis, let me point out, was not limited by any control for father's occupation. Subsequent work confirms these findings (see, in particular, Richard B. Freeman's *Black Elite* [1977]). The article by Siegel that Duster cites as counterevidence is 14 years old, based on even older data, and not relevant to this issue. Young blacks entering the labor force with higher education, Freeman shows, do as well as whites or better. Admittedly, the issue of higher black unemployment and older blacks who have suffered from discrimination remains.

Now, on the problem of policy, with which my book deals, there is no question that the economic position of any group is to be explained by large social and group factors. But it is a radical leap to go from this given to the conclusion that employment and educational policy should take into account the race and sex of individuals in making decisions on employment, promotion, admission, and graduation. It is social factors that explain why poor groups produce the best boxers. But if we were selecting boxers, would there not be an individual injustice, leaving aside the question of effectiveness, in ignoring this social reality in order to select equally from all groups? On the passage on the "incompetent and criminally inclined" that Duster quotes from my book, it is one thing to understand the social causes of incompetence and criminal behavior; it is a radically

different thing to argue (and it was in this context that I wrote as I did) that even though one group has a higher percentage of incompetents or criminals, these factors, because they are socially produced, should not be allowed to count in selecting persons for roles in which competence and noncriminality are generally agreed to be important.

I argue against the use of goals and quotas based on a statistical mirroring of the population as a policy to deal with what is admittedly a serious problem, the distribution of racial and ethnic groups in key positions, on many grounds: the difficulty of determining the line between what groups should benefit and what groups should be penalized, the difficulty of assigning individuals to given racial and ethnic categories, the political and moral repugnance at doing so, the danger that more and more groups would demand consideration in receiving benefits, and the danger that group membership would become ever more salient and group conflict increase. These still seem to me powerful arguments against our present course, even though we are still left with the question of what other policies we can find to deal with the problem.

NATHAN GLAZER

Harvard University

REFERENCES

Freeman, Richard B. 1973. "Changes in the Labor Market for Black Americans, 1948–1972." Pp. 67–120 in *Brookings Papers on Economic Activity,* no. 1. Washington, D.C.: Brookings Institution.

———. 1976. *Black Elite: The New Market for Highly Educated Black Americans.* New York: McGraw-Hill.

Hall, Robert E , and Richard A. Kasten. 1973. "The Relative Occupational Success of Blacks and Whites." Pp 781–98 in *Brookings Papers on Economic Activity,* no. 3. Washington, D.C.: Brookings Institution.

REPLY TO GLAZER

Nathan Glazer insists that the 1970s ushered in a new era in which it was "not possible to demonstrate" by recent econometric analyses "discrimination against young blacks [already] in the labor force."

First, he does admit to the apparently massive discrimination which keeps young blacks out of the labor force. Given his position about the importance of individual rights to employment based upon individual attributes, skills, and characteristics, can it be sufficient simply to admit to the issue of extraordinarily higher black unemployment and not bring this to bear upon his analysis? To wit, does systematic discrimination or do individual attributes account for the following?

> In 1976 the rate of unemployment was 7 percent for whites and 13 1 percent for blacks and other minorities. In August 1977 white joblessness declined to 6.1 percent, while minority unemployment increased to 14.8 percent.
>
> In 1971, when 15.1 percent of white teenagers were jobless, the unemployment rate for minority teenagers was 31.7 percent In 1976 white teenage unemployment stood at 18 percent, while 39.8 percent of minority teenagers were unemployed; and by August 1977 unemployment for minority teenagers had reached a staggering 40 percent.
>
> In 1974 the annual median family income for whites was $13,356, compared with $7 808 for blacks and $9,559 for Hispanics. For most of the past decades, the ratio of black to white family income has remained fairly constant while the dollar gap between the two groups continues to grow For example, in 1964 the median annual income for black families was $3,724 compared with $6,858 for whites. In 1974 the annual median family income for blacks increased to $7,808 compared with $13,356 for whites. While the ratio of black to white family income has remained fairly constant (at about 2 3), the dollar gap between the two groups has increased from $3,000 to $5,500.[1] (U.S. Commission on Civil Rights 1977, p. 1)

Second, econometrics takes the kind of data listed above and provides statistical controls and factoring to try to account for these differences. It tries to account for discrimination indirectly, by inference. In so doing, it misses massive direct discrimination. Even the most sophisticated econometrics is a fundamentally inadequate methodological technique for getting at racial discrimination when contrasted with research done in the natural setting, as witnessed by a recent study (July 1977) of the Department of Housing and Urban Development. In communities in 40 metropolitan areas, 300 blacks and 300 whites sought housing vacancies. After more than 3,200 audits, a clear picture of patterned, if sometimes subtle, discrimination emerged. In one of the more subtle forms, realtors "steer" whites into white areas and blacks into black areas, irrespective of expressed housing needs. Econometric analyses of secondary data miss even obvious cases of discrimination systematically. In the area of employment, the literature is replete with findings indicating how a significant proportion of jobs in the market never even get listed but are passed by word of mouth among friends and acquaintances. In such situations, there is no active, personal discrimination, but discrimination is certainly built into the structure of social relationships. Also, the rise of private employment agencies provides a mechanism for "steering" that functions like the realtors in the housing study.

An anthropologist or sociologist would surely predict that research in

[1] The figures quoted here are from the U.S. Department of Labor, Bureau of Labor Statistics, for the relevant years, fully cited and directly quoted.

the natural setting of personnel employment offices that review prospective employees would turn up results similar to those of the HUD housing study. Glazer, Freeman, Hall, and Kasten, by a simple change to a more rigorous methodological device, that is, more rigorous in its responsiveness to the relevant empirical phenomena under investigation, could then conclude that it is possible to demonstrate systematic discrimination against young blacks, even in the 1970s, even against those in the labor force.

Finally, on the question of social policy, as I have indicated elsewhere, there is no simple or single social policy that will handle the legacy of racial inequality that spans most of the nation's history. Affirmative action is hardly an effective remedy. Indeed, it is often misguided, frequently ridiculous in its requirements, usually a paperwork mirage. But despite the small difference it has made in the actual circumstances of blacks, it has unleashed a torrent of righteous indignation on the part of those who are ready to believe that we are in a new era where it is preferable to have no social policy for the redress of the worst enduring injustice of the republic.

TROY DUSTER

University of California, Berkeley

REFERENCE

U.S. Commission on Civil Rights. 1977. "Statement on Affirmative Action." Clearinghouse Publication 54. Washington, D.C : Government Printing Office.

COMMENT ON FOX'S REVIEW OF *PUNISHING CRIMINALS*

Vernon Fox, the reviewer of my *Punishing Criminals: Concerning a Very Old and Painful Question* (*AJS* 83 [September 1977]: 537–38), is entitled to his views but not to misrepresent mine. His misrepresentations go beyond tolerable limits. Let me give a few instances.

I am said to draw "heavily on the Old Testament" and to call on "Plato, Ovid, and other classical philosophers." Ovid is mentioned once; Plato once apropos of "anomie" and in the following two passages: "This assumption—which goes back to Plato—" (p. 16) and "the notion that to be morally wrong . . . is to be sick . . . is as old as Plato. . . ." (p. 123n.). The Old Testament is mentioned in a footnote, together with the New, to illustrate biblical views on vengeance. It is quoted once more in the epilogue. This is the extent to which I draw heavily on it.

Fox writes: "Citing the philosophies of Rousseau, Marx, Kant, and Mao, the author holds that bad conditions in society tend to corrupt 'natu-

rally' good men." (1) Mao is mentioned once (in the epilogue): "From Rousseau, to Marx, to Mao. . . ." (p. 263). (2) Kant is quoted six times with respect to deserts and not at all (contrary to Fox) with respect to "bad conditions in society" and " 'naturally' good men." (3) My point was that Rousseau et al. were *wrong* in holding that "bad conditions in society tend to corrupt 'naturally' good men." Fox attributes to me the view I explicitly rejected. (4) Fox asserts that I believe that "to search for causes implies mistakenly that the offender is a victim of society." (His meaning is intelligible though his syntax is not.) My point was, on the contrary, that causal explanation does *not* imply justification, though all too often it has been thought to do so. (5) I do not recommend "an increase of fines and compensation, as suggested in Deuteronomy and Exodus" as Fox believes. (Incidentally, neither of these sources suggests such an increase.) I did recommend "income day fines"; I also pointed out that restitution, though desirable, cannot be a part of punishment. I can find no connection with biblical sources or with increases.

It would serve no useful purpose to illustrate further Fox's deficient reading abilities. The impression he gives is that *Punishing Criminals* is an elaboration of biblical or classical philosophy. I regret to say the Bible plays no role in my book. I do deal with the philosophical justification of punishment in perhaps 60 of the 270 pages. I try to reconcile retributionist and utilitarian views and come down largely, but not altogether, on the utilitarian side. Mao, Marx, Plato, and the Bible are not involved. About 200 pages are devoted to recent empirical data and modern theories of crime and control and crime causation. Fox has not reviewed any part of *Punishing Criminals*.[1]

ERNEST VAN DEN HAAG

State University of New York at Albany

[1] Fox has not replied to this comment.—Ed.

Review Essay: A Sociologist Looks at Wittgenstein

Wittgenstein and Scientific Knowledge: A Sociological Perspective. By Derek L. Phillips. Totowa, N.J.: Rowman & Littlefield, 1977. Pp. xiv+248. $17.50.

Stephen Toulmin
University of Chicago

Like other antiphilosophers before him—Cratylus, Pascal, Nietzsche—Ludwig Wittgenstein (1889–1951) dreamed of "putting an end to" philosophy by showing the vacuousness and confusion embodied in the arguments and positions of traditional philosophers. This was perhaps too much to hope for. Instead, Wittgenstein's work has added fresh and volatile fuel to the very debate he despised, among both academic philosophers and others. One curious by-product of this debate has been a genre of books that seeks to analyze the implications of Wittgenstein's ideas and methods for the social sciences. In his well-known essay, *The Idea of a Social Science,* Peter Winch has construed Wittgenstein's views about language and explanation as closely akin to those of Wilhelm Dilthey. Meanwhile, Hanna Pitkin's *Wittgenstein and Justice* reads Wittgenstein's arguments as leading to a political theory somewhere between those of Hannah Arendt and Jacques Ellul. And now Derek Phillips, who edits *Theory and Society* from the Sociological Institute at the University of Amsterdam, has added an intriguing but frustrating book to this growing shelf.

The contents of *Wittgenstein and Scientific Knowledge* fall naturally into three parts. The two opening chapters give a summary account of Wittgenstein's life and thought—a formidable undertaking, since Phillips allows himself only some 50 small pages to do this. Granted the inherent impossibility of the task, the results are impressive. Phillips rightly emphasizes the moral passion underlying Wittgenstein's whole enterprise, although he does not fully show us the relevance of this passion to the actual arguments Wittgenstein presents. He also sets out in helpful form the basic contrasts between the *Tractatus Logico-Philosophicus* (1919) and the *Philosophical Investigations* (1952) and expounds the central notions ("language-games" and "forms of life," "understanding" and "obeying a rule") around which Wittgenstein's later arguments are constructed. In this first part, there are some felicitous touches, for example, Phillips's reference to Italo Calvino's *Cosmicomics* as exemplifying the kind of imaginative tinkering with our basic modes of life that alone can enable us to extricate ourselves—at least, momentarily—from our basic modes of thought. There are also some defects: for example, Phillips places too much reliance on W. W. Bartley's gossipy allegations, later withdrawn

for lack of substantiation, about Wittgenstein's supposed homosexual practices.

The middle five chapters of the book switch somewhat abruptly to the philosophy of science. Here, Phillips sets out an account of two "images" that have, in his view, dominated the subject in recent years: an ahistorical "image" associated with the logical empiricists and a more recent historical and sociological "image," of which Thomas Kuhn is the best-known representative. Both views (he tells us on p. 55) are "at odds with a more Wittgensteinian conception" of science, which it is his own aim to present and defend in the final two chapters. The dominant, positivist view is too much divorced from the actual business of "doing science": it presents science too abstractly, as "a *rational process* . . . governed by the timeless principles of logico-empirical methodology" (pp. 63–64). The newer view, in contrast, goes too far in the other direction and falls into the familiar problems of any "sociology of knowledge," notably, the problem of relativism. From Kuhn's point of view, how can we compare the "merits" of theories or ideas put forward by scientists in different "communities of science" or different epochs? Each community and period is free (it seems) to make up its own standards of judgment and apply them in its own favor.

As Phillips sees matters, these difficulties can be overcome by paying closer attention to Wittgenstein's central theses. Although our basic modes of thought may be incapable of any logical grounding (or "transcendental deduction," in Kant's words) they are nevertheless indispensable to us in a more pragmatic sense. For something that Phillips calls "the general language-game of everyday life" is specific to, and universal throughout, the human species, and this species-universal pattern of conduct "*has an epistemological and ontological primacy*" (pp. 86–89, italics in original). The intellectual claims of science need to be related back to this universal everyday language-game if we are to achieve the necessary "commensurability" between scientific theories and concepts that Kuhn's arguments apparently destroy. And it is only when we look to see how the language and practice of science are explained and discussed in their relation to the language and practice of everyday life—discursively, rather than formally—that we shall finally come to see how it is rooted in the practical conditions of our existence.

The argument of these central chapters is interesting but in certain respects perverse. Wittgenstein himself nowhere makes out any kind of case for believing in such a global entity as Phillips's "general everyday language-game"; still less does he concede that entity "epistemological and ontological primacy." On the contrary, everyday life and language were always discussed by Wittgenstein as involving great and varied congeries of "language-games," each of which needs to be reflected on separately; there was no suggestion in his arguments that they fit together into any single, overall pattern. (In the passage quoted by Phillips on p. 88, for instance, the definite article "the" refers to *one particular* language-game from these whole congeries, not to the *entirety* of everyday life

and language.) If Wittgenstein did not adopt any general position of the kind Phillips attributes to him, there were good reasons for this, for it would have meant answering a number of questions about "the natural history of the human species" that Wittgenstein was too knowing and too cautious to commit himself about. He did not argue that there *is* "agreement in concepts" between different communities of language users *because* there is "agreement in forms of life" between them. Rather, he argued that there *can be* such "agreement in concepts" only *to the extent that* there is in fact sufficient "agreement in forms of life." In what respects and to what extent is this, in fact, the case? That was not for him to say, nor was this, strictly speaking, a philosophical question at all. For his purposes, there was no need to take a psychological stand with, say, Jean Piaget and Lawrence Kohlberg against, say, L. S. Vygotsky and A. R. Luria, or vice versa. Just how far, and in what respects, the "forms of life" characteristic of different communities, cultures, and epochs do *in fact* share universal, species-specific features was something that could be allowed to "show itself" in due course as the outcome of psychological, anthropological, and historical researches.

That last question is—to be sure—a fascinating and highly significant one. In some respects, indeed, it is "the question of the hour," about which the rival views of Piaget, Luria, Noam Chomsky, Claude Lévi-Strauss, and a dozen others all have something to tell us. Derek Phillips is, accordingly, right to draw attention to its significance for the philosophy of natural science and social science also. Just to the extent that, and in those respects that, there are universal, species-specific forms of life (and, hence, language-games), people of different cultures and epochs will have a guaranteed *locus standi* when it comes to making cross-cultural and cross-historical comparisons about the merits of rival scientific ideas. (Giambattista Vico might say, "I told you so!") But there is no need to give Wittgenstein credit for framing that question, nor is there any virtue in basing one's answer on the authority of his reputation. If anything, calling him in to testify on a topic he took care to avoid during his own lifetime only confuses the issue.

Two briefer points, in conclusion. First, Phillips himself does not wholly escape from the traps against which he warns philosophers of science. Thus, he shares the misleading equation of the "rational" with the "logical" that has dogged philosophers since the time of Plato and so misses the clarification that comes from breaking with that equation. The task for us at the present time is not, as (for example) Paul Feyerabend teaches us, to see through the scientists' claims to "rationality"; instead, it is to reconstrue those claims as an insistence on open-mindedness rather than on some mythical "logico-empirical methodology." So the final, existentialist declaration with which Phillips' closes—that he is here presenting an account of the world of science "in a form in which it has meaning *for me*" (p. 222)—is a piece of needless rhetoric, as also is his insistence on the essential "groundlessness of our believing" (p. 221). It is only those who are still, at bottom, wedded to the Platonic claims of geometry

and formal logic that need strike such postures. The question of "grounds" is not one of framing "logical demonstrations"; rather, it is one of discovering where in our lives our beliefs are securely rooted. In that respect, none of Phillips's arguments do anything to prove our beliefs "groundless."

Second, if I may be permitted to pick a personal bone, Phillips repeatedly alludes to my work in blanket references to "people like Hanson, Feyerabend, Polanyi and Toulmin" (p. 64—mixed company!), and he even quotes at length my own arguments against Kuhn's relativism (p. 114). Yet he goes on to present as his own original insight a theory of scientific change that emphasizes "competition" and "survival" (pp. 160 ff. and 174 ff.), in terms barely distinguishable from the arguments of Donald Campbell and myself. I am more than happy to welcome Derek Phillips to the community of "right-thinking" philosophers of science. But it would have been civil of him to acknowledge that he has not invented that "evolutionary epistemology" totally unaided.

Review Essay: On Parsons

L'Individualisme institutionnel: Essai sur la sociologie de Talcott Parsons. By François Bourricaud. Paris: Presses Universitaires de France, 1977. Pp. 350. Fr 69.

Harry M. Johnson
University of Illinois at Urbana-Champaign

As François Bourricaud writes early in his splendid book, it has been the fate of Talcott Parsons to be widely misunderstood and attacked for reasons having little or no basis in his actual work. *L'Individualisme institutionnel* should help to move discussion to a new level, at least in France. It is easily the best secondary treatment of Parsons in French or English and, I suspect, in any language. Bourricaud self-consciously strives to treat Parsons in the same close and critical but uncarping spirit in which Parsons himself has approached the work of Durkheim and Weber.

Among the important topics in sociological theory that are examined closely are the paradigms of social interaction, of socialization, of deviance and social control, and of evolutionary change; the four-function paradigm; the pattern variables; the components and levels of social structure; and the generalized symbolic media of the social system (money, power, influence, and value commitments). Bourricaud, unlike most other commentators, gives considerable attention to the various meanings of "paradigm." In Parsons, there are several types; but the more important paradigms, such as the four-function scheme, consist, I think, of interrelated concepts referring to selected variable features of empirical reality. These concepts do not make up a description of any actual case but, rather, constitute an explicit framework within which many very different cases can be flexibly but systematically analyzed and to some extent explained. Moreover, the important paradigms named above are themselves interrelated. As Bourricaud writes, "they possess a structure if not identical, at least comparable" (p. 153). All these paradigms are as much concerned with process as with structure, with malintegration as with integration, with change as with stability. Bourricaud shows clearly and, one hopes, once and for all that Parsons has never remotely been a "functionalist" in the sense of that word that, sociologically, implies dogmatic belief in integration and perfectly consistent functioning and, ideologically, implies complacency about capitalism as the best of all possible worlds. Utilizing Parsons's theory and its implications, Bourricaud gives one of the best treatments of social change I have ever seen and deals at length with power, conflict, and force. His discussion of "pattern maintenance" and change, of the sense in which values may retain "identity" through time, and of the concept of "break-

through" is brilliant and subtle. Equally so is his balanced discussion of stratification, which interweaves both early and more recent treatments by Parsons. In general, he shows that symbolic systems, although important, work rather imperfectly and only under certain conditions (which he discusses). He shows that deviance, malintegration, conflict, and change are virtually inevitable: that is, theory makes their ubiquitous existence understandable and to some extent predictable.

One of the finest sections is a careful demolition of the idea that *homo parsoniensis* is "oversocialized." The internalization of culture, even combined with the pressures of the social situation and the physical and biological environment, leaves a real if limited scope for individual autonomy. The process of socialization, insofar as it is "successful," results not, to be sure, in "free will" in the theological sense, but in a greater "freedom" than the individual would otherwise have. Bourricaud calls this measure of autonomy "institutionalized individualism" and treats it as the result of socialization in any society. This is definitely in keeping with Parsons's ideas. Bourricaud also deals with institutionalized individualism in the more restricted sense of the relatively broad scope and high degree of individual autonomy that, especially after the Protestant Reformation, came to be institutionalized in some Western societies: what Durkheim calls "the cult of the individual."

It would take much more space than I have here to deal adequately with this rich book. Among its many merits are the following:

1. It brings out the continuity (with changes, of course) between Parsons's work and that of his predecessors: the Utilitarians, Marx, Durkheim, Weber, Freud (to name some of the more important). Like the Utilitarians, Parsons treats "action" as goal directed, and, like them, he gives a very important place (though a different one) to rationality. Like Marx, he recognizes that conflict is "structured," but of course his treatment of conflict is much less dogmatic, makes many more distinctions, and relates conflict to the normative order in a different way. The importance to Parsons of Durkheim, Weber, and Freud is probably better known, but Bourricaud's treatment, though necessarily selective, is characteristically intelligent and brings out both Parsons's indebtedness and his originality.

2. Since Bourricaud is obviously anxious to dispel widespread misunderstanding about Parsons and, of course, about theoretical and empirical issues, he interweaves a good deal of comment on Parsons's polemical orientations, explicit and implicit. Among the more important objects of polemical criticism are "functionalists," notably Malinowski (of whose theory Parsons's critique, I might say, is far more cogent than most of the criticisms directed against Parsons have been). Parsons has also consistently though quietly opposed the tendency toward cultural determinism found in some anthropologists (typified by Ruth Benedict). He has also consciously and consistently opposed "behaviorism." "Action" cannot be understood as simply *re*action to the environment or to the situation. *Homo parsoniensis*, as Bourricaud shows, makes choices and combines many kinds

of "information" to arrive at decisions that cannot be related to external stimuli on a one-to-one basis. These combinations-choices-decisions are guided, very imperfectly, by cultural codes, that is, by various kinds of symbolic paradigms. (Here we should note the difference between "guidance" and determination.)

Marx and some of his successors, such as the so-called critical theorists, have also, of course, been objects of Parsons's courteous negative criticism. This is a complex matter, which Bourricaud, it seems to me, handles very well. Following Parsons, he does not allow himself to get mired in the debates about the "real" Marx; rather, he deals with difficulties from which Marx never fully extricated himself at the level of coherent, explicit theory and in which many Marxists remain caught. Parsons's reasoned view is that, while prediction is extremely hazardous (not much is "inevitable"), both experience and theory make extremely unlikely ("utopian") a radical advent of "freedom" and equality, and the ending of anomie, alienation, "exploitation," and conflict. In this sense, Parsons as theorist is not a radical "integrationist" but rather, if you like, a "realist." At the same time, in a world not likely to become fully harmonious, he does not accept the Marxist critique of Western industrial society. Organized workers are far from powerless, even though their freedom (like that of owners, managers, and professionals) is far from unlimited. Moreover, society is not "produced" or "reproduced," in the ideological sense that Herbert Marcuse's work, for example, suggests. The image of puppets who are successfully manipulated unawares by a ruling class is a gross caricature based on crude theory. As for the Marxists or Marxist revisionists who concede that socialism might come into being through bourgeois democratic processes, their concession is fatal to a theory that has postulated a struggle to the death between two class groupings, a struggle in which, allegedly, the polity and the cultural system are largely instruments of the owning-exploiting class.

3. Bourricaud's book brings out, even though somewhat incidentally, the development of Parsons's own theory. At the same time, the reader is made conscious that the theory is being developed from basic starting points that have remained important throughout his work, although of course the successive phases do bring out new aspects and relations. Present from the beginning has been the actor/situation frame of reference, which Parsons distinguishes from the organism/environment frame of reference of biology. He recognized very early the relativity, abstractness, and interpenetration of "personality" and "society" and the importance for both of normative elements and of symbols in general. He recognized that "rational" action is possible and has meaning only within a nonrational (not irrational) framework. (At first he was more likely to speak of "ultimate ends"; later, of "values.")

As far as method or style of work is concerned, here too certain early patterns have been maintained. First, as we have noted, the strategy of building up and building upon analytical paradigms has been present from *The Structure of Social Action* (1937) on. Second, there is the constant effort to consolidate or integrate different writers and even different fields

by a careful analysis of resemblances, differences, and relations. Analogies, carefully and imaginatively examined and pursued, have played as important a part in the development of Parsons's theory as they have in the development of common law. A third feature of Parsons's style of work is his habit of focusing on problems—theoretical and empirical problems. For him, theory that has no relevance to empirical problems is not scientific theory at all.

Bourricaud uses the doctor-patient relationship whenever he can to illustrate theoretical points, but he by no means confines himself to this example. For instance, he devotes a long chapter to industrial society, discusses the American political system, and analyzes carefully the stability-through-change of the Christian belief and value system.

4. A fourth merit of Bourricaud's book has already been suggested but needs elaboration: his appreciation of the importance of "general" theory. He properly emphasizes the kind of theory that is possible and most likely to be useful in sociology. "In any case," he writes, "there is no reason to confine theory development to a descent from the general to the particular" (p. 155). Noting that Parsons has "modestly and serenely" declared that his theory is not a deductive system such as Homans favors, he writes, "For my part, I should be tempted to argue that in the sense in which Homans takes the word, there is no satisfactory theoretical system and, even more, there cannot be one, if we limit theory to pure deduction" (p. 155). Bourricaud then goes on to summarize the following passage from Parsons's essay "On Building Social System Theory" (*Daedalus* 99 [Fall 1970], p. 868):

> [My practice has] been a process much more like that of many developments within the common law. The work of which *The Structure of Social Action* was the outcome certainly established a theoretical orientation—in my sense of theoretical—which was not a congeries of random opinions. . . . From this . . reference base the process has been one of exploring a rather wide variety of highways and byways of empirical-theoretic problems, not, however, in wholly random succession. . . I have considered . . . topics and problems in relation to a theoretical scheme, which—though its premises were not defined with complete precision and henceforth assumed as fully given in a logically complete sense—has had considerable clarity, consistency, and continuity. In a sufficient proportion of cases, it seems to me that this kind of procedure has yielded empirical insight and rounding out, extension, and revision and generalization of the theoretical scheme. At certain points this has meant intensive concern with formally defined theoretical problems, but at other points primary concern with much more empirical issues.

Throughout, Bourricaud has obviously been conscious that he is dealing with a major theorist of the stature of Durkheim and Weber. Moreover, although Bourricaud is concerned with polemical issues, his tone is courteous and measured. He is not anxious to "put anyone down." For instance, he does not even mention the authors of some of the more hideously distorted treatments of Parsons's work.

Although I have called *L'Individualisme institutionnel* perhaps the best available secondary work on Parsons, I should emphasize now that its judicious balance and sense of proportion; its combination of exposition, evaluation, and direct and indirect treatment of criticisms; and, above all, its analysis of implications make it an original contribution as well, an exciting intellectual experience even for readers like me, who have already devoted a great deal of time to Parsons's work.

Book Reviews

Max Weber: A Biography. By Marianne Weber. Translated and edited by Harry Zohn. New York: John Wiley & Sons, 1975. Pp. xiv+719. $19.95.

David M. Trubek
University of Wisconsin—Madison

It is a great boon that Marianne Weber's touching biography of her husband is finally available in English. Originally written in the 1920s shortly after Max Weber's death, *Max Weber, ein Lebensbild* has only now become available to the English-speaking public in this excellent translation by Harry Zohn.

Marianne Weber attempted many things in this book. The book is not merely conventional biography: it contains extensive selections from Max Weber's letters to family friends and colleagues, charming memoirs of Max and Marianne's deep emotional and spiritual relationship, a vigorous defense of Weber's role in numerous academic and political controversies, and an introduction to the sociologist's major ideas.

Perhaps the last aspect of the book is its least successful or useful feature. In her explicit discussions of his theories, Marianne Weber tells us nothing new about her husband's contributions to social thought. This is not to suggest any lack of intellectual acuity on the author's part or any failure to understand Weber's complex thought. Quite the contrary, this remarkable woman, an author and active feminist in her own right, shows she was one of Weber's adept pupils. Her restatements of the themes of *Economy and Society* (Roth and Wittich, trans. [New York: Bedminster Press, 1968]) and the role of religion in the rise of capitalism are competent. But they tell us little that is not already available in the texts themselves and, because they so closely follow the texts, are not as useful as the many secondary analyses of Weber's ideas that have appeared in the years since this book was originally written.

While it offers little as a primer on Weber, the book is a moving document of great historical interest and high literary merit. Marianne Weber shows considerable artistry in the pictures she paints of Weber the public and private man; his family; and the social, political, and intellectual milieu of the Wilhelminian Germany in which he worked and lived. Moreover, in her description of Weber's struggles for ideas he fervently espoused, she establishes (at least indirectly) a vital link between the man and his ideas, thus contributing to a fuller understanding of the ideas themselves.

In the foreground of this superb portrait is Max Weber, the public person. We see him in his multiple roles first as a lawyer, then university professor, private scholar, and political activist. Almost as prominent are

the sketches of the private man: as a son torn by tensions between his autocratic father and his devout mother; as a husband whose devotion to Marianne sustained him in periods of great emotional difficulty; as a teacher tormented by the strains that academic life placed on his limited nervous energies. In the background are pictures of German university and political life at the end of the 19th century and into the troubled first decades of the 20th.

It has been said of Weber that he was at the center of most of the intellectual currents in Western social thought of his time, and certainly that is true. In addition, he managed to be open to a wide range of social and political developments. In *Max Weber: A Biography* we see him moving in the political circles of Berlin and leading major scholarly ventures while, at the same time, visiting with Stephan George and members of his mystical circle, spending summers in a colony of Bohemian freethinkers, meeting with radical and reactionary student groups, and, throughout it all, living a conventional bourgeois professorial life with Marianne in Heidelburg and Munich.

The narrative is selective, and Marianne does not deal with some of the darker currents of their lives. She does speak at great length about Weber's nervous breakdown which occurred shortly after a dramatic confrontation with his father. However, to a generation which has absorbed Freud and post-Freudian thought, her treatment of the causes of this protracted mental illness, which kept Weber from working for several years and from active teaching for more than a decade, seems curiously superficial. Further, though aware of Freud's ideas—widely discussed in their circle—Marianne is unwilling or unable to apply them to her husband's condition, which is presented as an inexplicable and mysterious disease. She discusses Weber's interest in eroticism, which emerged after the breakdown and became important in his later work, and relates discussions with friends about new forms of sexual morality without suggesting anything more than purely intellectual concerns. Where others have suggested that these intellectual concerns reflected Weber's personal experiences, Marianne uses this opportunity to assert Weber's continued commitment to traditional norms of Christian monogamy. Finally, she describes the death of Weber's sister Lili as an "accident," while Zohn corrects this statement (without comment) by noting that Lili took her own life (p. 687).

In these passages and in her discussion of the radical movements in thought, art, politics, and morality which impinged more and more on the Webers as the bourgeois culture of pre–World War I Germany cracked and then crumbled with the nation's defeat, Frau Weber strikes a characteristic tone. On the one hand, she wishes to be scrupulously honest in depicting these new ideas and the people who held them; on the other, she wishes to express the distance that separated her and her husband from them. While she expresses no moral revulsion at ideas that must have seemed strange and indeed threatening to the comfortable and pious German professional class from which she came, at the same time she strives to make clear that she did not accept, or choose to live by, these new ideas.

This effort to retain her own ideals, as well as a residual innocence and perhaps ignorance, may make it appear that Marianne has glossed over things that more "modern" scholars would wish to highlight. But at another level, this insistence on the maintenance of the ethical ideals that she and her husband shared emerges as the book's greatest strength and perhaps its most significant contribution to our understanding of her husband.

Marianne does stress Weber's ethical and political ideals, and her description of Max Weber's life emphasizes his struggle to realize them. Whatever omissions may result are not the biases of a loving spouse or the reflections of a residual prudery, but the incidental costs of the theme she has selected.

Once one realizes how clear sighted Marianne was, and how much she knew of the world of ideas and events in which she lived, it becomes apparent that her effort to see the world in terms of moral values is itself a moral effort. She describes the values she and Weber shared—dignity, individuality, devotion to others, duty. She realizes these values were constructed from an amoral universe through individual efforts. In her description, she not only shows her own exceptional personal strength, she also demonstrates a grasp of some of the most complex implications of her husband's thought. For this book deals, more than anything else, with the relationship between Weber's ideals and his actions. The picture that Marianne has painted, of a man ceaselessly struggling to realize his highest ideals in scholarship, teaching, politics, and private life, is a description of the human condition as her husband understood it.

While American sociology has tended to focus on Weber's demand for value-free science and for the divorce of scholarship and action, Marianne Weber wants us to understand the relationship between the life of the scholar and his values. Thus, in her discussion of Weber's ethical thought, she notes that he rejected any implication that his views on value-free sociology led to amoral relativism. He believed, she states, that "there can be different and equally justified opinions about the nature of moral obligations, but . . . it was an indubitable inner certainty that only the choice and recognition of ideals, tasks, and duties give meaning to human existence. It is our destiny to know that we cannot derive the meaning of what happens in the world from the results of our study of history, however good our research, but that we must be capable of creating this meaning ourselves" (p. 325).

The theme that unites the many incidents she recounts—from efforts to help friends threatened by bureaucratic stupidity, conventional morality, or political venality to grander struggles for political liberalization and academic freedom—is the heroic effort to create meaning by fidelity to ideals at all costs. One may not accept Weber's ideals and one may wonder whether Marianne has told the entire story in her recounting of these incidents, but it is impossible to doubt the overall veracity of her portrait or the nobility of her subject.

Nothing is more moving or more persuasive than the story she under-

stands best—their marriage. In one of her husband's most moving passages he spoke of the meaning of marriage as "the linkage of marriage with the thought of ethical responsibility for one another . . . can carry the sentiment that something unique and supreme might be embodied in marriage; that it might be the transformation of the feeling of a love which is conscious of responsibility throughout all the nuances of the organic life process, 'up to the pianissimo of old age,' and a mutual granting of oneself to another and the becoming indebted to each other. . . . Rarely does life grant such value in pure form. He to whom it is given may speak of fate's fortune and grace—not his own 'merit' " (H. H. Gerth and C. Wright Mills, eds., *From Max Weber: Essays in Sociology* [New York: Oxford University Press, 1958], p. 350). Anyone who reads this book will realize that Weber and his wife were so graced. The story of the Webers' marriage—of the mutual affection, respect, and need which was strengthened during the long crisis of his mental illness—is a story of "love conscious of responsibility." And once this is understood, it is possible to see precisely why this biography is a significant monument to both Max Weber the man and his deepest ideas.

L. T. Hobhouse, Sociologist. By John E. Owen. Columbus: Ohio State University Press, 1975. Pp. xii+225. $11.00.

Philip Abrams
University of Durham

The reputation of L. T. Hobhouse among sociologists is a curious one. The few who have made any study of him testify to his promise and importance. The majority, duly noting the fleeting significance of the fact that he was the first English professor of sociology and perhaps mildly surprised to see that he was not the sort of empiricist one might have expected an English sociologist to be, accord him a fairly marginal place in the background to the history of their discipline. John E. Owen's careful and lucid study of Hobhouse's writings does something to clarify this mystery.

Hobhouse, it seems, was a victim of his own intellectual milieu. A serious and conscientious man, he devoted his life to an attempt to achieve a synthesis of the social and moral philosophies of Hegel, Bentham, and Herbert Spencer. He sought a liberal theory of moral evolution governed by the idea of the progress of mind. Sociologists in general would, I suspect, like to believe in the possibility of such a project; more than anything else it would legitimate their professional existence and practice. We are, after all, in the knowledge-for-progress business. But the prospect is daunting, its philosophical hazards are grim, and it is comforting to dwell upon the failure rather than the challenge of the pioneers who tried to make the issue of the direction of sociology explicit. Along with Max Weber, Hobhouse is thus commonly judged to have attempted an impossible enterprise. If we

can see their life purposes as interesting but essentially confused we can sensibly lower our own sights to more mundane levels.

But was Hobhouse confused? Owen suggests that, on the contrary, he had a peculiarly clear understanding of the difficulties that beset the would-be liberal sociologist. How can one demonstrate that science advances progress while admitting that science rests on assumptions that are culturally relative? How can one ground the notion of progress in the actuality of human purposes in a way that makes development a meaningful idea but avoids teleology? How can one validate the claims of social science as knowledge while repudiating both idealism and positivism? To address such questions was to recognize that sociology exists in an uncomfortable space between philosophy and politics, and that if the sociologist wishes to speak to the politician he can do so effectively only to the extent that he has the blessing of the philosopher. Because Hobhouse saw sociology as a form of applied philosophy he insisted above all on the importance of the philosophical credentials of sociology. It was this worry about the intellectual basis of the subject as much as anything that inhibited him from actually doing empirical sociological work.

Sociologists who believe in getting on with the job have, as a result, always felt a certain impatience with Hobhouse. His rediscovery as a significant forerunner at a time when epistemological anxieties are again widespread is perhaps not surprising—especially because, as Owen convincingly shows, his attempt to master such anxieties was at least more substantial than the mere descriptions of the crisis which seem to be the most many sociologists can contemplate today. Within the terms of the philosophical tradition of his day Hobhouse does seem to have achieved a possible agenda for sociology, and it was not a trivial agenda. The sociology of morals was to provide an understanding of the conditions that governed moral progress—that is to say, of the conditions that would enable serious, decent liberals like Hobhouse to realize their purposes.

The trouble with Hobhouse, then, and this emerges very clearly from Owen's analysis, was not confusion but blandness. He lived in a world in which it was quite plausible to believe in the reasonableness, tolerance, and open-mindedness of others. In such a world, innocent of Hitler, Stalin, and the dialogue of moral absolutism that has characterized academic relationships since 1968, it was not wholly absurd to regard progress as a matter of informed argument. But even then it was a little whimsical and curiously remote. Even in the small sociological world of his own day Hobhouse was to find himself isolated and estranged from the majority of his colleagues. They thought him irrelevant, patronizing, and ineffectual because he seemed to have nothing to say about the practical issues of social policy that concerned them. But his larger irrelevance, his irrelevance for sociology today, springs from the assumption that informs all his work from *Morals in Evolution* to *Social Development:* that sociology is a discourse among liberals—that its intellectual problems can be confined within the limits of the assumptions and premises of liberalism.

It is a defeated as well as a gentle face that stares at us from the cover

of Owen's book. And the one thing the author does not explain is the expression captured by the photographer. As Owen guides us through Hobhouse's successive writings he documents the skill, delicacy, and coherence with which the last of the mainstream English liberal political philosophers developed the case for sociology as the science of the harmonious and progressive collective realization of individual purposes. But Owen's essay ends confidently, emphasizing the devotion of Hobhouse to his task and the erudition and sophistication he brought to it. Hobhouse himself constantly discovered that the world in which he actually lived and professed sociology was not at all the world of rational discourse, mutual understanding, and respect for philosophy postulated in the sociology he had chosen to profess. Rather, it was a chaos of conflicting realities lacking both the will and the means for the coordinated reasoned discussion which is the essential practice of liberal sociology. No wonder he looks hurt.

Philosophical Foundations of the Three Sociologies. By Ted Benton. London: Routledge & Kegan Paul, 1977. Pp. xi+225. $11.95.

Geoffrey Hawthorn
Cambridge University

The three sociologies in the title of this book are those modeled on what are taken to be the procedures of the natural sciences, on the hermeneutic interpretation of texts, and on the Marxism of Louis Althusser. Each, Ted Benton suggests, takes a different view of the proper relations between philosophical argument and claims about the world. For the first, philosophical arguments are preliminary and eventually inessential. They serve merely to clear the ground for a sound and self-sufficient science. In Locke's phrase, philosophers are underlaborers. For the second, philosophical arguments are intrinsic and essential. They constitute the phenomena about which claims are to be made and the grounds upon which these claims are to be secured. Philosophers are master scientists. For the third, "the methods and forms of explanation adopted by the sciences must be adequate to the distinctive character of their objects." This "involves rejection of the positivist external imposition of a universal scientific method without, at the same time, surrendering to the 'humanist' [i.e., hermeneutic] rejection of *any* notion of the unity of science" (p. 137). It requires "a set of systematically logically related concepts and propositions which is logically discontinuous with the conceptual frameworks in terms of which actors and groups understand themselves and their relationships," a set which "neither presupposes nor implies evaluative and normative commitments of a practical kind" (p. 127). It is a conception, at least one of the difficulties of which "is to give a characterisation of philosophy (and found a practice of philosophy) which avoids the error of setting up philosophy as the 'final arbiter' on all questions of cognitive status (the basic error of classical

theories of knowledge) without risking the disappearance of philosophy altogether (its disappearance into science, or into the history of science)" (p. 11). Benton sets out to defend this third sociology.

He starts by attacking the other two. No positivism, he argues, on the one hand, can on its own terms provide anything other than an instrumentalist or conventionalist account of causality or indeed of the phenomena to be explained. No consistent hermeneutics, on the other hand, can insist both upon the constitution of particular meanings by particular actors and upon the possibility of these meanings being externally caused and of thereby being instances of general explanations. These are fair points and familiar. Together they explain why, for all honest men, there are now no acceptable philosophical foundations for the social sciences. Benton, though, is torn between honesty and ambition. Accordingly, he sets off in pursuit of a decisive solution based on Althusser's Marxism, which he eventually suspects is no solution at all.

"If scientific knowledge is to be recognised as a reality," he insists, "distinct from ideological forms of knowledge, then we need a theory of the mechanism of its production" (p. 173). "This . . . is the only way of avoiding the humanist commitment to relativism or a philosophical anthropology without simultaneously resorting to positivism" (p. 217, n. 6). The question, of course, is, How to secure the theory to establish the theory which would show the theory to be true? Benton asks us to consider seriously Althusser's answer. Althusser himself asks us to suppose that a theory must be a theory of a real external world made up of the materially produced, to suppose that theory itself is materially produced, and yet to suppose that a theory " 'is indeed its own criterion and contains in itself definite protocols with which to *validate* the quality of its product, i.e., the criteria of the scientificity of the products of scientific practice' " (p. 188). That is to say, Althusser asks us to accept a most extreme thesis about the conditions of knowledge on the basis of a putative theory the validation of which owes nothing either to its conditions or to the appraisals of the theorist. He asks us to use our rational discretion to suspend belief in our rational discretion. Not surprisingly, neither he nor Benton is able to describe nonrhetorically the foundations Benton himself sets out to defend. Benton does wonder whether "the theory of science outlined here" is "a philosophical theory of knowledge, or . . . merely a theoretical approach to the history of science as a social practice for the production of knowledge" (p. 192). It is, of course, neither. It is an internally inconsistent assertion of a morally disagreeable faith.

Benton seems to be led to take it seriously for two reasons. One is his Marxism. The other is his commitment to "paradigm-neutral criteria of validity," to "progress in science" (p. 198), and, more generally, to epistemological security at any price. Together, these dispositions override his evident intelligence, precipitate him into preposterous nonsense, and commit him to exactly that conception of the metaphysician as master scientist which he begins by disavowing. Those who wish to concede a real world of real men, who believe that we should at least begin by taking these men

seriously, who are willing to entertain the fact that there are no philosophically sound grounds at all for being able to say that sociologists actually know what they are doing, and who are therefore prepared to accept that there are still no rules of sociological method, if by that is meant procedural rules for discovering the truth, those who are willing to see that precisely because of all this philosophical sophistication is a sine qua non of a sociology that is not vacuous, sterile, and insensitive, would do well to contrast Benton's book with John Dunn's "Practising History and Social Science on 'Realist' Assumptions" in *Action and Interpretation: Studies in the Philosophy of the Social Sciences*, edited by Philip Pettit (New York: Cambridge University Press, 1978).

Sociology, Phenomenology and Marxian Analysis: A Critical Discussion of the Theory and Practice of a Science of Society. By Barry Smart. London: Routledge & Kegan Paul, 1976. Pp. xiii+206. $16.00 (cloth); $13.25 (paper).

Robert W. Bogart
Illinois Institute of Technology

For some sociologists their discipline has become intellectually bankrupt, politically irrelevant, and morally deformed. These are harsh but genuinely felt sentiments. Such beliefs are inseparably bound to the more general conviction that the claim of sociology to have achieved the status of a scientific endeavor is completely fraudulent and illegitimate. Needless to say, beliefs and sentiments of this sort, and those they inspire in opposition to them, have generated controversies of immense proportions within the profession in recent years. What is perhaps more important, however, is that those most insistent in their denunciations of the discipline have been moved in response to their dissatisfactions to search for theoretical and methodological formulations that would offer distinct alternatives to those available in conventional sociology. The efflorescence of neo-Marxism and the elaboration of "phenomenological sociology" can be accounted for as an outcome of this search.

Barry Smart sets for himself the task of critically evaluating the many debates spawned by these contemporary developments. He is not, nor does he claim to be, an unbiased critic. He agrees with those adversaries and detractors of the profession who believe that conventional sociology is in a state of profound crisis. According to Smart, the source of the crisis is to be found in the fact that sociologists have unwisely based their endeavors on the questionable methodological cornerstones of a natural science model of inquiry. In short, sociology has allowed itself to become merely a "scientistic" undertaking. In so doing, it has lost consciousness of the intimate and inexpungeable linkage between theory and practice; it has, in fact, become atheoretical. It has unthinkingly adopted the assumptions of ana-

lytic positivism and in the process given total weight in its causal schemes to objective social structures thus completely overlooking the intersubjective structures which actually serve to sustain the reality of everyday social life.

Smart does not believe that neo-Marxism has fared much better because it too finds itself committed to the same set of unexplicated methodological premises intrinsic to "scientism" in general. This, Smart claims, is not the fault of Marx. Rather, the fault lies with his credulous proponents who have chosen to disregard the methodological critique of scientism which Marx adumbrated in his various works, especially those centrally concerned with exposing the implicit reified nature of the categories of analysis dominant in classical political economy. By a similar argument, the sociological alternative originating in existential phenomenology (e.g., ethnomethodology, phenomenological sociology, reflexive sociology, etc.) turns out, upon closer scrutiny, not to represent in its present form a significant alternative at all. While these more exotic approaches to sociology have been brilliantly critical of scientism, the substance of their prime theoretical concerns has led them away from the most vital questions. Instead of trying to explain the forces at work shaping the way individuals live in society, the major thrust has been toward explaining how individuals talk about and conceptually formulate their experiences. This represents an essentially ahistorical and ultimately reductionistic perspective because, at base, even linguistic patterns and conceptual categories are preconditioned by the recurrent and objective forces at work in society.

The obvious question—Where does all this leave us?—then becomes paramount. Smart proposes, and I am simplifying his argument a great deal, that the only truly acceptable alternative involves an approach which draws together Marxian analysis and phenomenological sociology. Such a reconstituted phenomenological Marxism would greatly facilitate a fundamental redirection of sociological inquiry. It would allow, and indeed require, that the "dialectical" relationship between the objective and subjective processes on which social life is based be taken into full account. It would force the sociologist to renounce positivism and the naturalistic model of analysis embodied in the physical sciences. And, perhaps most important, a phenomenological Marxism would restore an interest in creating a truly critical sociology, a sociology concerned with enlivening and canalizing the emancipatory instincts of modern man.

For anyone who has been relatively attentive to the course of intellectual events of the past decade, there can be no escaping the conclusion that the discipline has not made any demonstrable progress toward resolving the controversies dealt with in *Sociology, Phenomenology and Marxian Analysis* since their original expression in, for example, the remarkably profound work of Max Weber in the early part of this century. The reasons for this lack of progress are not difficult to understand, and though largely insignificant otherwise, *Sociology, Phenomenology and Marxian Analysis* at least serves to exemplify these reasons.

It is full of obscure jargon. No matter how many times one uses such

terms as "scientism," "positivism," and so forth, they remain incomprehensible until they are assigned an explicit meaning. However, like most of his colleagues, Smart eschews formulating his categories of criticism in terms of an understandable, and thus publicly available, language. Moreover, irrespective of one's assessment of the relevance of the issues treated in this and similar works, the task of clarification cannot be assisted so long as those engaged in such efforts are unprepared or unwilling to expend as much time and intellectual energy in the constructive as in the merely critical aspects of their undertaking. Since sociology is primarily a research discipline, the ultimate utility of any particular theoretical or methodological approach is calculable only in terms of the contributions it makes to our understanding. If phenomenological Marxism possesses any worthwhile explanatory insights, then it is incumbent upon those so convinced to demonstrate the fact by research and not by assertion alone.

Another, and possibly more crucial, weakness of this and other works in the same genre is that they are less concerned with truth than with impressing their political intentions upon the discipline. Discovery of truth can be possible only if it is clearly recognized that a statement of fact and a statement of value preference are entirely different and separable phenomena. Since in this work Smart is concerned with obliterating this most fundamental distinction, I can only conclude that his arguments lack any warrant whatsoever as a legitimate criticism of anything. The list of the insurmountable difficulties that are intrinsic to endeavors of the present sort could be considerably extended. Need more be said about their value?

On Critical Theory. Edited by John O'Neill. New York: Seabury Press, 1976. Pp. 263. $12.95.

Russell Jacoby
University of California, Los Angeles

Even those completely sympathetic to the "critical theory" of the Frankfurt School will find themselves impatient with this volume. *On Critical Theory* is a collection of 11, apparently new, essays. I say "apparently new" because the editor, John O'Neill, does not state that they are new (nor in fact does he mention them). No previous publishing history is provided although at least one essay—to my knowledge—has appeared elsewhere.

The reasons why O'Neill does not mention the essays may lie close at hand: The essays defy summary and unity and generally contradict O'Neill's introductory comments. O'Neill suggests that Habermas's "linguistic turn" risks renouncing the older and more powerful critical theory tradition of Adorno, Horkheimer, and Marcuse and that Habermas's emphasis on the normative function of rational discourse is losing the heart, soul, and memory of critical theory. Moreover O'Neill has elsewhere (*Sociology as a Skin Trade*, New York: Harper & Row, 1972) soundly criticized

Habermas. Yet with the partial exception of one essay, no one here takes up O'Neill's ideas; in fact the reverse is true. Most of the essays are close in spirit and ethos to the most recent writings of Habermas. This is unfortunate, both because it deprives the book of any inner unity and because a collection of essays critical of Habermas from this perspective would be useful.

The Habermasians are represented most emphatically by Albrecht Wellmer ("Communication and Emancipation") in a sympathetic discussion of Habermas's linguistic "reconstruction" of historical materialism. Friedrich W. Sixel ("The Problem of Sense") provides an exposition of the current German debate between Habermas and Niklas Luhmann. Dieter Migelt explores and criticizes Habermas's reading of Hans-Georg Gadamer ("Critical Theory and Hermeneutics"). This essay, as well as some of the others, suffers from the peculiar jargon that seems to inflict the Habermasian debate. Undistorted communication "requires that one risk one's preunderstanding as an act of practical discourse, and that one engage in a communicative action denying, in its accomplishment, its own communal basis" (p. 182). So much for undistorted communication. Regardless of style, these three essays would be useful to those unfamiliar with Habermas's more recent endeavors.

The essay by H. T. Wilson is the only one consistent with O'Neill's introduction. In "Science, Critique and Criticism" he discusses Habermas, Popper, Weber, and Marcuse. Habermas's defense of Weber against Marcuse's critique is evaluated to the disfavor of Habermas, who is seen capitulating to a Parsonian scheme. Wilson writes that "recent developments in his thinking clearly portend the end of critical theory" (p. 226). This is a concise and intelligent discussion of the issues and may be the best essay —perhaps because I agree—in the book.

Two more essays add to the Habermasian atmosphere. Here it is not a question of a presentation of Habermas but an application. Jeremy J. Shapiro seeks to work out a "new evolutionary theory of embeddedness in nature and its transcendence" (p. 156). Embeddedness in nature (*Naturwüchsigkeit*) is a concept traced back to Marx. Shierry M. Weber suggests a model of praxis based on art and self-reflection—very thin stuff.

Other essays include: Paul Piccone's over-confident denunciation of the Frankfurt School as hopelessly obsolete, and Christian Lenhardt's serious reflections on Horkheimer and Adorno's *Dialectic of Enlightenment*. Two more essays are not last but least. One could be characterized as a Yippie attack on Adorno, appealing to John "wild sociology" O'Neill and Herbert "new science" Marcuse as the new gurus. Ben Agger tells us, in the book's opening essay, that "between Auschwitz and Haight-Ashbury critical theory has changed its tune" (p. 32). Now theory rocks out, makes love, and Adorno is a bummer from the past. The clinker in this collection, with some very tough competition, may be Ken O'Brien's defense of Erich Fromm. This essay has serious problems with English, German, concepts, and facts. While O'Brien parades his knowledge of German, the result is a disaster. For example, he gives us four variant spellings of Fromm's title "Zum

Gefühl der Ohnmacht," none correct. (If no one proofread the book, perhaps part of the guilt lies elsewhere.) His acumen for facts can be judged by the following: referring to Marcuse's (and my) critique of Fromm published by Beacon Press he notes: "It is ironic that both the somersaults and social histrionics—the allegations that Fromm is revisionist and therefore [?] pretheoretical and bourgeois—are carried out by left writers under the aegis of what are obviously pages of multi-national, corporate, capitalist publishing media" (p. 127). Now while few would dispute that the means here are pages, Beacon Press is a small independent firm, rather smaller than the publisher of this volume.

In brief, this book limps and lacks the virtues of a good collection. Too many of the essays should not have been published. Moreover, no focus is established, or, perhaps, even attempted. On the one hand, a book exclusively critical of Habermas according to O'Neill's introduction or even concerned with the whole of Habermas—early and more recent—might be useful. But this book is neither. On the other hand, as a presentation of the multifaceted concerns of contemporary critical theory, the collection is far too limited—and too Habermasian. The Frankfurt School has inspired —especially in Germany—material on fascism, aesthetics, political economy, and psychoanalysis. No hint of these dimensions has entered this collection. There is a danger in the Anglo-American world that critical theory is seen as ending—literally or metaphorically—with Habermas. O'Neill, despite himself, has ill served critical theory.

The Sociological Theory of C. Wright Mills. By Joseph A. Scimecca. Port Washington, N.Y.: Kennikat Press, 1977. Pp. v+148. $9.95.

Charles Derber
Brandeis University

The notable feature of this study of the work of C. Wright Mills is the effort to delineate a consistent theoretical model which not only underlies the full range of Mills's writings but also represents an important alternative to functionalism, conflict theory, and contemporary versions of Marxism. The proposition that Mills's contribution extends beyond a series of brilliant but intellectually eclectic critiques to the development of an innovative, theoretically coherent paradigm of radical sociology is argued forcefully here and will be of interest not only to radical sociologists—particularly those who do not find conflict theory or Marxism to be adequate intellectual frameworks—but also to others concerned with the development of contemporary sociological theory. Unfortunately, Scimecca has done little more than sketch the bare outline of Mills's model. His study lacks the theoretical depth and stylistic verve of Mills himself. Moreover, the substantive significance of Scimecca's work as a contribution to radical sociology is seriously diminished by its failure to probe fully the relation of Mills

to Marx and to explore the possibility of a creative synthesis between Mills's social-psychological framework and the Marxist paradigm of social structure and social change.

The core of Mills's model, as understood by Scimecca, stems from an integration of the social-psychology of pragmatists such as Mead, Dewey, and James with the historically and institutionally oriented sociology of Weber and Mannheim. Mills was seeking a social theory encompassing a fully human concept of the individual and a fully developed concept of social structure. In pragmatism and Mead's social-psychology, Mills found the roots of a socially based theory of selfhood and social character which preserved the element of spontaneity and volitional action so central to his conception of the human spirit. From the Weberian tradition, Mills derived a sophisticated, historically based concept of social structure and institutional orders. The synthesis of these frameworks, systematically elaborated in *Character and Social Structure,* was the basis of Mills's sweeping framework which integrated the structural and historical features of society with the most intimate features of human personality. As a paradigm for radical sociology, Scimecca argues that Mills's working model, while incorporating Marxist premises regarding the centrality of historical and institutional analyses in social theory, advances beyond Marxism in its rich social psychology and its refined nondeterministic structural analysis and is the foundation for a fundamentally original and distinctive radical tradition.

The usefulness and viability of Mills's framework for radical sociologists and, more generally, for contemporary social theory depends, however, not on its development as a paradigm distinct from Marxism but rather on its synthesis with a broadly humanistic Marxist tradition. Mills's model, while inspiring a brilliantly creative new focus on the interplay of institutions and social personality, offers by itself no comprehensive theory or perspective on the historical development of social systems and no systematic model of social change. Indeed, his model of character and social structure, stressing the molding of persons by institutions and the integration of personality by social role, in certain respects is perilously close to Parsonian and functionalist models of social system analysis that are characterized not only by an underdeveloped theory of dynamics but also by a model of man who—if not oversocialized—maintains the most precarious capacity for self-determinative behavior.

Mills's vacillation in his thinking regarding the historical agency of social change, shifting from the working class to the intelligentsia to the young largely in response to new events rather than theoretical considerations, testifies further to the weakness of his model regarding historical dynamics and social change per se. Yet if Mills's social-psychological paradigm is deficient precisely where Marxist theory is most developed, the strength of his work is in an analysis of social character and consciousness largely compatible with Marxism but either neglected or vulgarized by Marxist theorists. Theory directed toward the integration of Mills's model with the Marxist tradition suggests the possibility of a sophisticated neo-Marxist social psychology in which historical, structural, and psycholog-

ical constructs are fully interrelated and in which the vision shared by Marx and Mills of a theory bridging objective institutional realities and subjective domains of consciousness might be realized. A fully developed social psychology of social classes; a refined analysis of bureaucratic culture and consciousness; a model of personality, interpersonal relations, and everyday life under capitalist culture and institutions; and the analysis of the relations of the economic, political, and military spheres under different historical epochs of capitalism are all projects central to a Mills-Marx synthesis. Indeed, current developments in Marxist thinking, including a new working-class theory and the theory of the relative autonomy of the state, suggest lines of convergence already in progress. *The Sociological Theory of C. Wright Mills* should alert Marxist thinkers to other new directions, particularly in the realm of social character and social psychology, but is not likely to inspire a distinctive vision of radical sociology uniquely grounded in Mills's intellectual framework.

The Concept of Cultural Systems: A Key to Understanding Tribes and Nations. By Leslie A. White. New York: Columbia University Press, 1975. Pp. xiii+192. $10.00.

George Becker
Vanderbilt University

In the "Editorial Note" to *The Concept of Cultural Systems* we are told that this, Leslie White's last completed work, represents his "most mature thinking on the nature of, and the science of, culture" (p. vii). For those, however, familiar with White's other theoretical statements (especially *The Science of Culture* [New York: Farrar, Straus & Giroux, 1949] and *The Evolution of Culture* [New York: McGraw-Hill, 1959]), his "latest understanding" of culturology should prove a disappointment in at least two respects.

To start, entire sections of this volume are merely a reiteration of arguments encountered in White's earlier writings. Accordingly, we find explicated once again the development of the evolutionary perspective in the social sciences. This is done, ostensibly, to demonstrate that during recent centuries scientific theory "has been groping toward a concept of culture and a new science: culturology" (p. 1). What recent intellectual trends reveal, White argues, is that Western society has evolved sufficiently to embrace the premise that culture constitutes "the most powerful of all determinants of human behavior" (p. 125). Since all sciences are the product of an evolutionary development, he continues, a fully developed deterministic science of culture has been made possible only during the 20th century. As in former writings, White's objective continues to be, in his words, "to make the trends of the past more explicit [and] to organize the various conceptions dealing with culture into a coherent theory" (p. xii). The pre-

occupation with this objective appears essentially self-serving. It is designed to legitimate the author's contribution as constituting the only logical and evolutionary culmination of preceding speculations on culture. My criticism is not of White's continuing objective per se but of the practice of providing a series of formerly encountered arguments in virtually unaltered form—arguments that remain tied to many of his earlier references and illustrations. The degree to which White relies on his former observations can be seen in his making reference to or citing from a total of 18 of his own publications. In addition, few other works are cited that were written after 1960. Of a total of 189 sources contained in the bibliography, only 21 were written after 1960, and three of these are by White.

The value of White's contribution, however, must ultimately be sought in those portions of the volume that propound a supposedly innovative approach to an understanding of preliterate and modern societies. It is the concept of cultural systems and the corresponding method of vector analysis that are advanced as the only suitable tools for the scientific study of society. Even though there cannot exist a culture apart from man, White argues, culture must be considered without reference to the human species, since all culture traits act and react among themselves in accordance with the principle of cause and effect. The culturologist, accordingly, must focus on what are identified as the three fundamental components or classes of elements of all cultural systems: the technological ("tools, weapons, and techniques of using them"); the sociological ("customs," "institutions," "codes"); and the ideological ("ideas," "beliefs"). As in William Ogburn's distinction between "material" and "adaptive" cultures (see *Social Change with Respect to Culture and Original Nature* [New York: B. W. Huebsch, 1922]), it is the technological element that is defined as the "great prime mover" (p. 178), since it imposes change and directionality on the sociological and ideological elements. Each of the three cultural components, in turn, is seen as comprising a large variety of "structures"—organizations, groups, and interests. The technological component, for example, consists of structures such as agriculture, mining, and manufacturing. Since each of the structures within the three cultural components represents specific interests that are measurable in terms such as size of membership, quantity of goods produced, or acres of land cultivated, they are defined as analogous to mathematical vectors—they possess the properties of direction (or "objective") and "magnitude." Cultural vectors, in this view, have the force to influence other vectors and the cultural system as a whole. Each strives to have its own interests served regardless of the welfare of other vectors or the larger system. Out of this competition of interests emerges a new synthesis reflected in a corresponding course of action which leads to either change or the retention of the existing order. Given these properties of cultural vectors, White concludes that cultural phenomena are subject to the method of vector analysis. It is with this method that the author proposes to understand the behavior of "tribes and nations."

It is not feasible here to evaluate critically White's numerous culturological and evolutionary premises which "dictate" vector analysis. It is pos-

sible, however, to assess the value or potential contribution of vector analysis. We must concern ourselves with a central question: What insights does this approach afford that would not be revealed by other existing means? It is against this question that White's work must be judged a disappointment. Let me briefly demonstrate.

White employs his theory of vector analysis to explain a wide variety of phenomena. His examination encompasses such far-ranging topics as the development of exogamous/endogamous regulations, the origins of social stratification, preferences in art, the lack of gun control legislation, the energy crisis, the American involvement in Vietnam, air pollution, the "irrationality" of political systems, and the futures of mankind and the nation state. The fact that White manages to treat these and many other topics within the confines of 180 pages is a reflection of the shallowness of vector analysis. For example, White asks, Why it is that cultural systems engage in behavior that is often contrary to their best interests? Pointing to supersonic transport (SST), he observes that a rational course of action would have dictated the development of less dangerous and cheaper forms of mass transportation. The reason the SST was heavily subsidized by governments and became a reality—in the face of a desperate need for mass transit systems in most metropolitan areas—can be comprehended, says White, only in terms of cultural vectors. He reminds us that the aircraft industry, like every other vector, is concerned only with pursuing its own interests. In the world of vectors "it is everyone for himself and let the devil take the hindmost" (p. 80). In the contest among the different technological vectors for support and government funding, therefore, it was the "aircraft vector" that possessed the greatest force to influence other vectors and the cultural system as a whole. It is in this sense that "cultural systems are not ethical or moral systems. . . . They behave in terms of their intrinsic properties in accordance with the principle of cause and effect" (p. 88). Their behavior "is the expression of the integration of its components or vectors" (p. 91).

White's analysis with regard to his other areas of concern proves equally unsophisticated and reductionist in nature. We are told that the absence of federal gun control legislation—despite the fact that 70% of Americans appear to favor it—is due to the greater organization and power of the antilegislation vector headed by the National Rifle Association. The reason Western art, unlike that of China and Japan, has emphasized the nude human form is, similarly, the consequence of a dominant vector. White observes that in art "one could choose—or hit upon—this or that subject and treat it in this or that way. But once a choice had been made . . . it would develop along the path adopted and would exclude other ways and forms: the original vector would inhibit the development of other and different vectors" (pp. 62–63). The same explanation, he maintains, is applicable to preferences in musical styles and the Judeo-Christian preoccupation with the "inherent sinfulness of the human race and salvation through the Redeemer" (p. 63).

In conclusion, White's work is little more than a restatement of his

former evolutionist and culturological position. The introduction of the concepts of cultural systems and vector analysis fails to enhance this original position in a meaningful way. One has to wonder, in fact, whether his use of vector analysis is indeed capable of providing anything beyond a crude reductionism—one that explains the existence of phenomena as that which must exist.

The Painted Word. By Tom Wolfe. New York: Bantam Books, 1976. Pp. 120. $1.75 (paper).

Albert James Bergesen
University of Arizona

Obviously Tom Wolfe is on to something. But no one seems to acknowledge what he has seen. The critical reviews of *The Painted Word* all seem to go on about his superficiality and lack of knowledge about modern art. Considering the fun he makes of two very sacred institutions, modern art and New York intellectuals, I suppose it is no wonder that the critics should reply in kind. But the significance of *The Painted Word,* at least for sociologists, lies not so much in Wolfe's moral evaluation of recent trends in art or in his descriptions of the various social types which make up the art world as in his insightful observation of two trends in modern art: the gradual de-objectivization of art—its becoming more and more abstract, and the corresponding rise in the necessity of some kind of art theory. These two points are worth pursuing for they suggest ways of understanding change in artistic style, and more generally, a way of conceptualizing art as a kind of linguistic code.

In discussing the history of modern art from early modernism through abstract expressionism, pop, op, minimal, earth art, and conceptual art, Wolfe observed the more or less general trend of removing more and more content from the art object itself. "In the beginning we got rid of nineteenth-century storybook realism. Then we got rid of representational objects. Then we got rid of the third dimension altogether and got really flat (Abstract Expressionism). Then we got rid of airiness, brushstrokes, most of the paint, and the last viruses of drawing and complicated designs (Hard Edge, Color Field, Washington School)" (pp. 97–98). He points out how we next got rid of the picture frame altogether (Frank Stella's shaped canvases). Then we got rid of the picture altogether and painted on the walls of the gallery (Robert Hunter and Sol Lewitt). Then we got rid of the walls themselves, creating sculptures which filled the museum (Carl André, Robert Morris, Ronald Bladen, and Michael Steiner). Then we got rid of the gallery and museum with earth art, such as Robert Smithson's *Spiral Jetty* in the Great Salt Lake. And finally, with conceptual art, we got rid of the idea of a permanent work of art itself, so that, ". . . there, at last, it was! No more realism, no more representational objects, no more

lines, colors, forms, and contours, no more pigments, no more brushstrokes. . . . Art made its final flight, climbed higher and higher in an ever-decreasing tighter-turning spiral until . . . it disappeared up its own fundamental aperture . . . and came out the other side as Art Theory! . . . Art Theory pure and simple, words on a page, literature undefiled by vision . . . late twentieth-century Modern Art was about to fulfill its destiny, which was: to become nothing less than Literature pure and simple" (pp. 107–109). While I do not feel entirely comfortable with this sort of broad sweeping account of the changes in modern art, I do feel that there is enough truth to his point about increasing abstraction to provide some crude evidence for a conception of art as language and for extending our understanding of changes in sign systems in general.

What Wolfe has captured here is the complex relationship between the vocabulary and syntax of an art object, as expressed in questions of color, form, line, brushstrokes, and so on, and the larger cultural context, art theory, in which the art object is embedded. Wolfe has described what Basil Bernstein refers to as elaborated and restricted linguistic codes. Bernstein argues that the meaning of a linguistic act (or in this case an artistic act) comes in part from the actual internal substance and structure of the speech act (vocabulary and syntax) and partly from the more general set of understandings the group holds as to the meaning of these linguistic signs. What Bernstein calls restricted codes develop in situations in which there is a smaller lexical pool and a narrower range of syntactical alternatives from which to construct a speech act. These codes arise in groups where there is a high degree of solidarity, where people know each other on a face-to-face basis, and where there are many commonly held assumptions so that there is little need to spell things out. Hence the restricted code. Elaborated codes, on the other hand, arise in social situations where there is less solidarity, less face-to-face contact, and accordingly fewer commonly held assumptions. Here meaning derives from the internal structure and substance of the language as found in the complexity and variety of vocabulary and syntactical arrangement. Since people do not know what you are about, you have to spell out your intentions. Hence the code is more elaborate. There is an implicit equation here, what I will call the semantic equation. On the one side is the linguistic act, or art object itself, and on the other side are the assumptions the group takes for granted. The dynamic part of this equation centers on the fact that changes on one side of the equation are reflected in changes on the other. If the pool of vocabulary and syntactical alternatives decreases, then the broad cultural understandings must increase so that the semantic equation can remain in some sort of equilibrium, that is, the sentence or art object can still be a meaningful act to some audience.

How does all this apply to Wolfe and modern art? Modern art—in becoming increasingly abstract—throwing out line, color, brushstrokes, and picture frames—is reducing its artistic vocabulary and syntax and, as a means of communication, is becoming an increasingly restricted code. Since the origin of meaning for an art piece is found decreasingly within the

actualities of the painting itself, there has to be an increase in art theory to give the art object meaning. The semantic equation remains balanced. As content declines, theory increases. In effect, if a person says less through a speech act or a painting, you have to know more about what he is talking or painting about to make sense out of it. This reliance on the broader cultural context, here art theory, to make sense out of a restricted code is nicely captured in a observation of Hilton Kramer's: "Frankly, these days, without a theory to go with it, I can't see a painting."

How restricted can a code become before a crisis in the semantic equation is reached? In the case of modern art the continual removal of elements from the art object can seemingly reach a point where all the meaning in the semantic equation resides on the side of the cultural context, giving us object free art, as in conceptual art where we have art without any sort of visual experience. "Art . . . disappeared up its own fundamental aperture . . . and came out the other side as Art Theory! . . . Art Theory pure and simple, words on a page, literature undefiled by vision" (p. 109). Now the emergence of a new realism in the 70s, such as what Wolfe terms "photo-realism" in his epilog, suggests that there may be a cyclical character to these changes in the semantic equation. As a code becomes more and more restricted, as visual free art is reached, so to speak, the semantic equation is no longer balanced because the cultural context now carries virtually all of the meaning. This kind of system is unbalanced, collapses, and then begins again. The new realism, therefore, may be the product of the completion of a cycle of ever restricting codes, which become so restricted that they disappear, with the whole semantic machinery resetting itself in some sort of new equilibrium.

There is still the question of why the semantic equation should move in one direction or another. The world of New York art, with the close contact and familiarity of artists, critics, intellectuals, gallery owners, patrons, and museum directors represents the same sort of intimate and closed society that Bernstein found in his working-class neighborhoods. To the extent that the New York art world was this sort of close intimate community, it seems reasonable that it developed a restricted code in its artistic expression—for example, the various sorts of abstract, minimal, and conceptual art. The decline of New York's role as an intellectual and cultural center represents a weakening of that community's sense of centrality and, as such, a movement away from artistic communication through restricted codes like abstract and minimal art to more elaborated codes, such as the photo-realism mentioned by Wolfe.

All of this, of course, is highly speculative. But we do know that at the height of New York's fame as the center of modern art in the 40s and 50s, we had the emergence of abstract expressionism as a highly restricted code and the increasing necessity of art theory (Clement Greenberg and Harold Rosenberg) through which abstract art was interpreted, and more fundamentally, made into something that could meaningfully be understood as "art." The decline of New York's importance is also associated with the collapse of the secret language, the rise of the new realism—photo-realism

—and the increased attention being given to photography. As the community withers, its commonly held assumptions similarly dissipate and the burden for meaning falls on the other side of the semantic equation, that is, on the system of signs rather than on the larger cultural context of art theory. What is to be artistically spoken can no longer be hinted at, or activated by an abstract clue or gesture. It must be said in full by the coherent ordering of the vocabulary and syntax of the art object itself. In some sense, I suppose, the appearance of Wolfe's essay is also part of this process, for with the larger cultural context which gave meaning to abstract forms dissolving, the abstract forms now stand somewhat naked, without a larger purpose and meaning, and, as such, open to ridicule.

Music and Society since 1815. By Henry Raynor. New York: Schocken Books, 1976. Pp. viii+213. $15.00.

Stephen Hart
Wesleyan University

This is a social history of Western music since the Congress of Vienna. It describes changes in the social and political environment of music, related changes in the economic and organizational aspects of musical life, and the manner in which all these changes have conditioned the work of composers and performers. It is mainly concerned with the period before 1900, with developments in the major European musical centers, and with "formal" or "serious" music. Within these boundaries, *Music and Society since 1815* is a thorough and informative account which should be of use to anyone concerned with the sociology of modern Western music.

For convenience we might distinguish among three ways of approaching music sociologically. The first starts from the fact that music, like any other social activity, requires economic resources, physical facilities, and appropriate skills and attitudes, and that it is often embodied in formal organizations, a distinctive set of professions, and so on. Clearly such factors condition the actual creation and communication of music. A second approach deals with the social functions of music, crudely as an adjunct to war, dancing, religious ritual, work, or courtship and more deeply as a powerful force in the creation of solidarity in many social contexts. The third, or cultural approach, treats music as a symbolic structure, like language or religion, which has its own inner coherence and patterns of evolution. Henry Raynor, though he mostly eschews conceptualization, operates primarily within the first of these approaches, and therein lie the contributions and limitations of this book.

Raynor starts with the decline of court patronage, which meant that to an increasing extent starting near the end of the 18th century composers and performers had to survive economically largely by selling

tickets to audiences. Composers sought out new audiences; virtuosi marketed their charisma; conductors, impresarios, and orchestras struggled to make a living by responding to the demands of a rising bourgeoisie. In general, making profits (or minimizing the deficit if an enterprise was subsidized, as were most operas and orchestras) became a critical determinant of what kinds of music were created and performed. New forms of organization evolved through a process of structural differentiation: the roles of composer, conductor, and administrator—united in the 18th-century *Kapellmeister*—became distinct. Orchestras grew not only in size but also in specialization, and by the 19th century needed conductors to coordinate their diverse functions. The social and psychological functions of music also became specialized. A bifurcation developed between a new "light" orchestral music created (in Raynor's view) to provide pure entertainment (and, through simplification and sentimentalization, to make music correspond to the lowest common denominator of taste) and serious music with lofty aesthetic purposes which became less and less accessible to any mass audience. This last differentiation, unlike the other ones, Raynor sees rather negatively as a sign, result, and possibly a cause of a "fragmented culture" in which there are few overarching bases for solidarity, few common cultural experiences. In addition, Raynor traces some of the political forces influencing musical life: censorship (especially in Austria and Italy) reflecting Metternichean reaction; nationalistic aspirations; the hope cherished by many that music could replace religion as a source of humane values; and educational policies which, according to Raynor, inhibited communication between composers and audiences apparently by stifling the development of the skills needed to comprehend and respond to new music.

This analysis, especially where it concerns the problems musicians faced in making a living and the strategies by which they survived, is detailed, well documented, geographically comprehensive (in most chapters, Raynor analyzes successively developments in Austria, various parts of Germany, France, and England; he then devotes a special chapter to developments in the United States), and often fascinating. In case we have forgotten, it reminds us both that musicians have the same economic needs as other people and can choose only among that range of alternative activities which will allow them to subsist and that musical organizations are subject to the same constraints and tendencies as any others. One should note, not necessarily as a criticism, that the range of issues characteristic of the second and third sociological approaches to music (as I outlined them above) are not much dealt with. In particular, little attention is given to the nature and causes of changes in musical style—a question which I would argue should be treated sociologically, not just left in the hands of critics and musicologists.

Raynor, writing in the idiom of historians, is reluctant to place his data theoretically. Sociologists can apply sociological labels (for instance, structural differentiation) which may allow them to relate what Raynor describes to other processes occurring at the same time. Occasionally, the

absence of theory is frustrating. For instance, many of the developments Raynor describes appear to fit well into a Marxist conception of "bourgeois music." That is, the criterion for the production of music becomes a money balance—there is a whole chapter on the commercialization of opera. Virtuosity becomes a commodity hawked like any other (see e.g., Raynor's discussion of Paganini on p. 56). Very significantly, amateur music spreads in a commercialized, privatized form; it provides a large amount of support for professionals (e.g., through selling music and giving lessons) and increasingly takes to the home. With the advent of the piano one can make music by oneself instead of with other musicians, and with broadcasting and sound recording one need not leave home even to hear professionals. Accompanying these structural changes is, at least in the 19th century, an increasingly subjectivist style of composition and the idea that music should express highly individual, private feelings. Furthermore, music becomes thought of as an autonomous enterprise parallel to law, politics, science, and economic life. One of the frequent characteristics of formal music, I think, is the conception among its practitioners that their enterprise serves no purposes other than its own—that it is pure entertainment or pure art. While this notion is not new, its dominance is, for historically most music has accompanied and supported nonmusical activities—marching to war, performing tedious work, dancing and socializing, or celebrating religious occasions—and has been answerable to nonmusical criteria. The point of these observations is to argue that developments in music might be clarified by a more explicit theoretical perspective.

This is, as I have mentioned, a book about what could be called formal or serious music—music written down and usually performed by specialists for socially elite audiences, typically on occasions which are specifically musical. (This is only a very rough definition, as there is considerable dispute about what distinguishes serious music from other musical forms.) Obviously limitations of scope are valid in historical work, but the choice of this particular focus does cause a problem which should be mentioned. Considerable confusion surrounds Raynor's discussion of the differentiation of serious and popular music, and of the "fragmented" situation Raynor sees now. In one important passage, he comments:

> While the composer was left to depend upon his ability to find the audience for whom his music was intended, social and political circumstances made it difficult for him to do so. The educational policies fostered in the period of repression—from 1814 to 1848 and for some considerable time, in some states, afterwards—deliberately cut as many of the means of communication as possible between the intelligentsia and the common people. The man of culture had no direct access to communication with a working class brought up, according to the dictates of Francis I [of Austria], to be simply good, honest citizens who knew their place and were contented with it. The composer found himself confronted by a social system which demanded that he found his audience but denied him the facilities for doing so. [P. 32]

According to this passage, the education of the working class deliberately stifled its competence to serve as an audience for new music. As a result, composers became hostile to the working class (as well as to political authorities) and either could not or did not try to communicate with it. Their music (see pp. 147 ff.) grew ever more intimidatingly complex, making excessive demands on both performers and audiences, and creating a gap which came to be filled by specialists in popular or light music.

One might raise arguments about the causal model found in this line of argument. However, there is even a prior question: the existence of the phenomenon to be explained. As Raynor puts it: "The development of light music towards ever-increasing simplification and of serious music into increasing abstraction and specialisation has created a schism so that it is often easy to see the two as acrimonious armies engaged in hostilities across the divide" (p. 151). It may be true that composers in the 17th and 18th centuries wrote, on various occasions but for the same kind of audience, both "heavy" and "light" music, whereas in the 19th and 20th centuries many (not all) composers have specialized in one or the other. But to evaluate whether this manifests an increasing fragmentation of music, a failure as Raynor sees it of composers to communicate with the masses, one would have to take into account changes in technology and economic life which have altered the boundaries of what kinds of music are visible to historical analysis. This is a methodological problem parallel to that of comparing the economic activity of societies with different proportions of the economy monetized. It seems quite possible that German peasants of the 18th century, instead of assembling to listen to the works of Bach (except perhaps in church), may have participated in musical activities which never resulted in written compositions, printed handbills, and concert programs, let alone sound recordings. The music of contemporary Western working classes does have such material results, and so is more apparent to us, and makes us conscious of the stratified character of our music. Furthermore, given rapid economic growth, the rise of urban classes with discretionary cash income, and the changes Raynor describes in the economic structure of music, attendance at concerts probably included a broader social spectrum in the 19th century than before, and thus one would expect a broader array of musical styles to be expressed in these events. Given these methodological dilemmas, Raynor does not provide sufficient information to prove his case.

Furthermore, when Raynor argues that because only a small segment of the concert-going public is able to respond to avant-garde music, composers have abandoned the goal of communication with audiences, one thinks of the alternative interpretation: that given the technology, economics, and sheer massive scale of the contemporary music scene, it is feasible for a large variety of audiences to exist, each small in percentage but significant in absolute numbers. Thus the evidence given does not establish the lack of an intention to communicate or a failure to do so. This question, again, could only be resolved by considering evidence of a dif-

ferent kind than Raynor provides. These issues however, while important for the polemics implicit in Raynor's book (and of course, the doubts I raise reflect in part my own musical feelings), do not significantly impair the usefulness of the book as a whole.

The American Dimension: Cultural Myths and Social Realities. Edited by W. Arens and S. Montague. Port Washington, N.Y.: Alfred Publishing Company, 1976. Pp. xvii+221. $5.95 (paper).

John MacAloon
University of Chicago

Anthropologists thumbing through the catalog of their 1974 convention came upon a symposium intriguingly titled "North of the Border: Anthropological Inquiries into American Culture and Society." The general reaction was wary amusement: "wary" because the failure to bring insight won in "the bush" to the systematic study of American life is a *scandalum magnum* of the discipline; "amusement" because the contributors had chosen to tackle the problem with a certain *esprit de jeu*.

The list of titles included: "The Bagel: Symbol and Ritual at the Breakfast Table," "Moonshining in Illinois," and "Good Marx for the Anthropologist: Structure and Antistructure in 'Duck Soup'." Other papers considered soap operas, astrology, professional football, adolescent literature, volunteer firemen, rock concerts, *The Exorcist,* poker, and the American fascination with the coffee cup. For those who missed the fun in Mexico City, the symposium is published in the present volume edited by W. Arens and S. Montague. Additional essays include the most eyebrow-lifting title of all, P. Claus's "A Structuralist Appreciation of 'Star Trek'."

In the introduction, some of the factors which have inhibited the anthropological investigation of our own society are noted: its scale and forbidding complexity; the problems of distance and detachment; and professional priorities which, until very recently, have insisted upon field work in "exotic" societies for admission to the guild. While recognizing the traditional wisdom underlying these constraints, Arens and Montague protest that the discipline has been hamstrung by them. Anthropology has been prevented from delivering the knowledge most worth having—the knowledge of ourselves. With this judgment one must concur. Unfortunately, *The American Dimension* does more to renew, than to overcome, professional reticence and caution.

The papers are diverse in conception and quality, but almost all of them are haunted by at least one of several related difficulties. This is first of all the issue of methodology. Face-to-face observation and questioning have been the hallmark of ethnography and the lifeblood of anthropology. Yet in the papers concerned with the symbolic analysis of entertainment genres

created for and transmitted by the media, there is little evidence of controlled, face-to-face interaction between the observers and the producers/consumers of these cultural products.

All will agree that the media "provide a gold mine of useful information for the analysis of American culture" (p. ix). But most anthropologists and sociologists will be taken aback by some simplistic assumptions stated nakedly in the "Introduction" and recurring implicitly in several of the papers. Because movies, television, pulp literature, etc., depend on a mass audience, can one automatically assume that "what is presented reflects the ideas and interests of a large segment of the American populace" (p. vii)? Will those who retain empirical aspirations and allegiances rest easy with the claim that, through the "cultural approach," "it becomes possible to directly ascertain [i.e., without bothering about native exegesis] the content of basic cultural symbols" (p. xvi)? Media content *is* "far easier to obtain than data collected by interviewing many individuals" (p. vii), but "ease" is hardly the criterion on which methodological decisions ought to be made. The vexing problems of research design which accompany field studies in pluralistic societies are here evaded rather than solved.

Some interesting interpretations are offered in this group of papers. Notable among them are I. Karp's suggestion that behind the humorous personae of the Marx Brothers lie the socially salient and marginal figures of the con man, the immigrant, and the androgynous clown; J. Kirkpatrick's claim that the "dramas of competence" and "dramas of loyalty" (p. 74) in situation comedies and commercials reflect alternative solutions to the conflict between individuals and their roles in American nuclear families; and S. Bean's excavation of competing categories of blood tie and love from the plot lines of soap operas.

But the reader longs for some extended case study which would put these interpretations to the test. In the absence of careful investigation into who (besides the authors) are receiving what messages from whom, the papers fail to persuade that the thematic or content analysis of cultural products actually reveals culture patterns (much less the social realities promised in the title). Moreover, little is done to overcome the skeptic's persistent tendency to regard work in one's own culture as moonlighting.

Another set of questions surrounds the "right relations" between established theory and new data drawn from novel sociocultural milieus. Karp and Arens feel the need to convince us—the former in passing, the latter, tortuously—that moviegoing and football are quasireligious rituals in order to justify their anthropological interest. Must new wine be forced into old bottles? The anthropological analysis of ritual is fulsome, whereas attention is just now being turned to ludic and aesthetic behavior and performance. Novel topics accrue legitimacy through analyses which generate new concepts and models and not through theoretical violence worked upon the data. The anthropology of literate, industrial society will gain legitimacy when it feeds as well as feeds upon work carried out in traditional contexts.

A related problem crops up in the two most arresting papers in the book. Claus applies Lévi-Strauss to "Star Trek" and interprets an episode of the television fantasy as the master interprets myth. It is not my desire to quibble with French structuralism here, but it does seem to me that Claus passes too rapidly over the multilayered metaphor embodied in the "Starship Enterprise" (military unit, corporation, nation, humanity, human body [p. 26]). These metaphors and the values and social circumstances of the program's young American devotees may have more to do with its symbolic power than the "orthodox" oppositions and mediations Claus hastens to discover.

S. Regelson's wild but alluring interpretation of bagels, lox, and cream cheese draws its inspiration from the studies of E. Leach and M. Douglas on food taboos. Regelson insists that the Sunday eating of these "significant substances" is a "ritual" linked "through a series of intermediate metaphors" to a "symbolic reversal" of dietary laws, incest prohibitions, and attitudes toward the millennium (pp. 136–37). These findings will be as controversial as the paradigms upon which they are based. But the point here is that Regelson does not pursue either the ethnic and religious values and circumstances of American Jews which he sees expressed, or why his informants must be so utterly "unconscious" of all this.

A final criticism concerns the duty to make oneself aware of the findings of other disciplines. Arriving in America is not like disembarking in New Guinea. Sociologists, psychologists, and economists have been here for some time. (So too, by the way, have anthropologists like Lloyd Warner, who gets a notice in two or three bibliographies but whose work is otherwise ignored.) In this respect, D. Jacobson's paper on labeling theory and friendship among middle-class engineers and, to a lesser extent, R. Jones's essay on poker playing succeed where others in the book fail. Both are solid contributions because they incorporate field studies and are truly alive to social realities.

If I have dwelt upon the shortcomings of this volume, it is because I share the contributors' conviction that cultural anthropology has much to offer to the understanding of American society, and that, in turn, general anthropology will be enriched. While *The American Dimension* will not attract a large following among professional anthropologists and sociologists, it is already serving some other important audiences. It has been used to good effect in high school and junior college courses to promote student awareness of the subliminal content of their own popular entertainments and microsocial practices. Moreover, I am informed by a school teacher on the Northern Cheyenne Reservation that his students delightedly find in its pages evidence for how "strange and foreign" the *wihio* really are.

The Puritan Way of Death: A Study in Religion, Culture, and Social Change. By David E. Stannard. New York: Oxford University Press, 1977. Pp. xii+236. $11.95.

John Carroll
La Trobe University, Australia

The adjacent territories that belong to history and to sociology have been under repeated invasion of late from both sides of the border. It is not that genuine differences between the disciplines have been broken down. Rather, many historians have been attracted to the theoretical tools and methods of sociology, and in return sociologists have reminded themselves that the disposition of the present inherited its genes from the past.

The book under review was written by a historian trying to extend the boundaries of his topic in the direction of sociology. It is a worthy book and yet it provides a cautionary example of multidisciplinary endeavor. This review is written purely from the sociological side of the border since I am a sociologist contributing to a sociological journal that is read presumably in the main by sociologists. I shall not be concerned with the accuracy, completeness, or representativeness of the historical evidence on which *The Puritan Way of Death* is based.

David E. Stannard tells a good story, and he tells it by and large in a suggestive and convincing manner. He traces changing attitudes to death among New England Puritans from the time of first settlement to the late 18th century. To illustrate his account he draws upon written theology, sermons, diaries, descriptions of burial rituals, and the iconography of headstones. As he proceeds he interprets the story of cultural change that he is telling, referring to a background of changing social forces in New England.

The main thread of the story is as follows. Medieval Catholicism had believed that the prayers and activities of the living could serve to shorten a dead person's term of purgatorial suffering. The Reformation reversed this view through its stress on predestination (in particular, Calvin's assertion that the fate of a soul is decided before death); it thereby removed the need for funeral ritual. However, from the middle of the 17th century, New England Puritans diverged from the austerity of their Old England brethren, becoming less and less restrained in their funeral practices. Stannard argues that the growing elaborateness of burial rituals was due to mounting insecurity about the legitimacy of the New World settlement. He takes up V. Gordon Childe's thesis that funeral rites become more extravagant during periods of social and cultural instability. The death of an individual, especially a leading citizen, becomes a symbol of the threatening disintegration of the society as a whole. This trend in New England was reversed after 1740 when the Romantic view of death as release started to take over. Cemeteries became neglected. Laurel wreathes carved on headstones were used to crown symbols of the victory of the soul;

a century earlier they had crowned the skeleton visage of death himself.

Stannard also examines the ambivalent attitude of 17th-century Puritans to death. Death was both longed for as the ultimate release and feared as an emblem of the fall. One of the strange consequences of this ambivalence was that Puritans tended to keep an aloof distance from their children. There was a great likelihood of a child's early death from disease; distance cushioned parents from the shock. But equally, according to Stannard, distance protected them from spiritual contamination: Children were likely as not "polluted with sin and natural depravity" and therefore not worthy of excessive affection. Such coolness on the part of parents would have caused their children overwhelming anxiety about rejection, an anxiety that Stannard plausibly suggests manifested itself in a morbid fear of death. Such fear was indeed common among Puritan children.

The Puritan Way of Death may be a significant contribution to the history of New England Puritanism. However, as a contribution to the sociology of culture it is flawed. The main problem is that it hardly ever incorporates the theory that it introduces. On one page (p. 92), for example, the concepts of anomie, badly explicated, and cognitive dissonance are introduced, and so is Hegel on progress. Yet nothing in this crowd of disparate concepts is used effectively. Stannard seems to have felt the need to tack his descriptions on to bits of theory in the hope that he might end up with a completed jigsaw puzzle, showing a vivid, clearly defined, and harmonious landscape. He repeatedly introduces suggestive theory (from Clifford Geertz and others) but fails to position himself in the middle of the tension between that theory and his own empirical findings. It is precisely out of the discomfort of living with this particular tension that innovations are made in social theory.

The first half of the chapter "Death and Burial" is highly effective—an elegant and lucid account of theological change leading to changes in social practice. But in the latter part of the chapter, as in much of the rest of the book, summary theories are introduced from the outside and the argument starts to jerk, to double back on itself, and generally to lose its bearings. Fortunately, the fascinating material that Stannard has discovered continues to speak for itself. *The Puritan Way of Death* would have been much better told as a straight historical tale without sociological trimmings. Moreover, the story is compelling enough for the reader himself to generalize it to other cultural situations. It is a pity that the author was tempted into visiting some of social theory's new cities, much like a nervous tourist, when he could have stayed at home and stuck to what he was good at.

Identity and the Sacred: A Sketch for a New Social-Scientific Theory of Religion. By Hans Mol. New York: Free Press, 1977. Pp. xvi+326. $13.95.

Robert Towler
University of Leeds

As we are told by the subtitle of his book, Hans Mol has set out in this volume to give us a sketch for a new social scientific theory of religion. This is a bold endeavor and as worked out it rests on two broad premises. The first is the very definition of religion with which he works, namely, the "sacralization of identity." By identity he means a sense of place in the world, and the search for identity, he says, "may culminate in drug addiction for some adolescents in the 1970s, or in the Jesus Freak movements for others, [but] is also recognizable in the most primitive of tribes who projected it on the totemic animal" (p. 2). The idea of a search is an integral part of Mol's perspective, for the model he presents is an essentially dynamic one. Religion is not this or that definable set of beliefs and practices. He wants nothing to do with substantive definitions of religion, though he recognizes their utility for certain limited purposes, but insists that religion is rather the process of an identity being sacralized, whether the identity be defined in terms of "celibacy, sex, war, death, beauty, the sacred, the profane, reason, feelings, science, equality, women, nationalism, virginity, buffaloes, snakes, tea, peyote, alcohol, action, tranquillity" (p. 2) or whatever (in this list and elsewhere he follows Klapp's *Collective Search for Identity* [New York: Holt, Rinehart, & Winston, 1969]). By sacralization Mol means "the process by means of which on the level of symbol-systems certain patterns acquire the same taken-for-granted, stable, eternal, quality which on the level of instinctive behaviour was acquired by the consolidation and stabilization of new genetic materials" (p. 5).

The idea of consolidation and stabilization links with the other main premise of his endeavor. He presents a high-level model of society as containing two countervailing forces which generally he refers to as differentiation and integration. These are the balanced centrifugal and centripetal impulses of yang and yin or of the adaptive and the integrative prerequisites in Parsons's scheme which are in a constant dialectical relationship. And religion is defined here at the level of culture: "Religion, or whatever sacralizes identity, has been put squarely on the stabilizing side of the dialectic" (p. 262).

These two themes form the backbone of *Identity and the Sacred*. The discussion ranges very widely but either the material is used to illustrate the two central themes or else the themes are employed to throw light on the material under discussion, and the author never strays far away from these main ideas. The book is extremely stimulating and well informed, and Mol succeeds in weaving together these twin motifs into a perspective of his own. But does it, as the author maintains, amount to a sketch for a new social scientific theory of religion? In my view it does not. The author

nowhere claims originality for either of his principal themes but neither does he say just how explicitly each of them has been spelled out by writers before him. The first represents no substantial advance on the thesis put forward by Thomas Luckmann in his *Invisible Religion* (New York: Macmillan, 1967), though the terminology is different and there is greater emphasis on sacralization as a process. The second is the argument set out by Victor Turner in *The Ritual Process* (Ithaca, N.Y.: Cornell University Press, 1977). However, in his discussion of Turner (pp. 240–43), Mol concentrates on *communitas* and antistructure to the exclusion of Turner's twin notions of *societas* and structure. And this is but one example of the main fault of Mol's book. In seeking to advance his own thesis he does justice to none of the writers he discusses. He takes from each what is needed to support his own perspective and criticizes the remainder for failing to fit into it. Turner's fault is that he does not identify religion solely with the "conservative" side of the dialectic, emphasizing instead its disruptive role in antistructure on a par with its integrating function in structure. But Turner is not alone: Durkheim, Weber, Berger, and many others are treated in the same way. There are exceptions. Swanson is not faulted, but his argument is distorted and many distinctions which are important to him are glossed over. Marx is found to have said nothing of value.

A general theory must either accommodate existing theories or show them to be erroneous and supersede them. Mol does neither. Because his new theory cannot accommodate the old ones he is forced to attack, but in so doing he fails to destroy and merely distorts them and sells them short. I would suggest that the time has not come for a unified theory, and so what Mol has given us is another perspective to rank beside those we have already, no two of which were compatible anyway. "Concentrating as heavily as we have done on the sacralization process contains the risk of a conservative bias" (p. 262)—that sums up well enough the flavor of Mol's particular perspective in this provocative and thoroughly worthwhile book.

A Sociology of Belief. By James T. Borhek and Richard F. Curtis. New York: John Wiley & Sons, 1975. Pp. xiv+201. $9.50.

Gary D. Bouma
Dalhousie University

In *A Sociology of Belief* James T. Borhek and Richard F. Curtis analyze belief as an important feature permeating all areas of social life. Their goal is to develop a set of strictly sociological, as opposed to social psychological or phenomenological, propositions about belief systems and their interrelationship with social organization. Their treatment is original rather

than a review of past research or a rehash of grand theory, and it is scholarly rather than popular or introspective.

This theoretical framework treats belief systems as supraindividual phenomena requiring supraindividual explanation. After delineating the elements and variable properties of belief systems, Borhek and Curtis explore the interrelationships between belief systems and social organization.

The basic elements composing belief systems are classified under the following rubrics: values; criteria of validity (statements about the grounds and methods by which statements are judged to be acceptable); logic (rules governing how to relate one substantive belief to another); perspective (a cognitive map of the relevant space); substantive beliefs; prescriptions and proscriptions (norms, policies, and action alternatives); and technology (statements about the way things are to be done). The fact that this list of rubrics is only a list is one of the weaknesses of this book. The categories are neither mutually exclusive nor are they analytically related. Values, prescriptions and proscriptions, and substantive beliefs seem to be the basic elements of a belief system while criteria of validity and logic seem to be a subset of prescriptive and proscriptive statements regarding the elements that are admitted to the system and interrelated within it. Technological statements appear to be norms, and the examples listed under perspective include all three basic types. At minimum a two-layered classification—basic elements and basic congeries of the basic elements —would be more satisfactory.

Borhek and Curtis classify belief systems by the following abstract variable characteristics: (1) The degree to which the elements are systematically related; (2) the degree to which the substantive beliefs in the system are empirically relevant; (3) the degree to which the system is open to change and innovation; (4) the degree to which the belief system tolerates other belief systems; (5) the degree of commitment required; and (6) the style of internal organization of beliefs. Six styles of organization are presented but again they are ad hoc rather than analytically distinct. It might have been more useful had the authors more deliberately and systematically imported insights from the general systems perspective in developing all of their categories.

Having described what belief systems are, Borhek and Curtis develop propositions about the relation between belief systems and social organization. From the "axiom" that belief systems are cultural, they derive propositions about patterning, change, differentiation, and institutionalization of belief systems. From the axiom that belief systems are grounded in both a context of social organization and a context of meaning, they argue that the two basic problems facing belief systems are maintaining commitment and validating beliefs. The continuing vitality of a belief system depends on two things: the survival of the group which has adopted it and the continued adherence to it by the group. Continued adherence depends on the social processes by which the beliefs are validated. This argument is muddied somewhat by the unsystematic inclusion of less directly relevant material, such as strategies of group survival. The section on

the social processes by which beliefs are validated is the strongest contribution of the book. In it the abstract variable characteristics of belief systems are most faithfully interrelated with the social processes by which belief systems are validated. Finally, while belief systems are dependent on a group for continued survival, the internal logic of a belief system may lead that group in directions not originally foreseen. In the closing chapters the general propositions are applied to the problem of belief in urban society.

In summary, this is a very useful contribution to the systematic sociological treatment of the phenomenon of belief. It is free of the institutional dependency so common in the sociology of religion and of the unproductive debunking approach so characteristic of the sociology of knowledge. It would have been much more worthwhile had a more deliberate attempt been made to import the general systems framework. The authors might have confused levels of analysis less frequently, seen the interrelationships between belief systems and social organization more clearly, and developed more analytically useful categories. Nonetheless, this is a very welcome and refreshing attempt to do some rigorous theory building in an area that needs it badly.

Ideology and Social Order. By Eric Carlton. London: Routledge & Kegan Paul, 1977. Pp. 320. $18.00.

Morton G. Wenger
University of Louisville

The scope of this work is both historically and analytically ambitious. It attempts an understanding of the relationship between ideology and social order which rests on a historical investigation of the nature of the ancient Egyptian and classical Athenian societies. In the process of developing his argument, Eric Carlton identifies his major goal as the demonstration of the "independent" social force of ideology and hopes to provide a means of typifying societies which is not based on stratificatory phenomena (e.g., caste, estate, class). Unfortunately, this interesting piece of social history fails at the level of sociological analysis, not so much because of its own endogenous defects as because of the faults of the general perspective within which it exists—Parsonian functionalism.

In a sense, this treatise represents a reworking of Weber's far superior *The Protestant Ethic and the Spirit of Capitalism* in a different and comparative historical context. Curiously, although Carlton cites this work in his bibliography and pays extensive methodological homage to Weber, he presents no specific discussion of *The Protestant Ethic*. This is unfortunate since Weber's cautious stance on the reciprocal relationship between religious ideology and economic organization would have served Carlton well had he emulated it. *Ideology and Social Order* takes a far more radi-

cal idealist position on the role of belief in society than Weber ever did, and thus this work is concomitantly more subject to the type of criticism directed toward Weber in Tawney's *Religion and the Rise of Capitalism*. Even if Carlton had adopted Weber's stance, a number of additional and probably insurmountable problems would have remained.

First and fundamental among these defects is one the author himself identifies as characteristic of cultural deterministic models—their "irritating circularity" (p. 238). Carlton aptly typifies his own argument when he states that "it is a sociological truism that the organizational patterns of societies are related to the control mechanisms which operate to maintain their structural viability" (p. 237). It is this truism that is belabored for 13 dense and lengthy chapters. Carlton attempts to extend this discovery to the point where the nature of control mechanisms takes on causal force in the structuring of society as a whole. The question that is scrupulously avoided throughout this book is *why* certain types of control mechanisms come into being. *Why* does Athens develop a "seminal" society while ancient Egypt is "archaic"? Carlton alleges that the effectiveness of social control depends on the type of ideology a society develops. Why does a society develop a particular ideology? This is the question that is not addressed, and it is the central one underlying the phenomena Carlton studies.

Given the equation of the relationship between social control and ideology, the above questions can also be asked in a more fundamental and revealing way: Why is *social control* necessary? It is in answering this that the ideological stance of *Ideology and Social Order* is revealed. Social control is seen not as a requirement brought forth by social inequality but, in the classical conservative sense, as a given—a prerequisite for the orderly functioning of essentially unruly individuals. Conflict is treated as infrasocial and not historical in nature. This is consonant with the general functionalist position on the value and function of inequality as well as on the relationship between individuals and society. This perspective leads Carlton to enter the matrix of argument at a point where the essential issues have already been decided, that is, the origin of ideology.

Had Carlton looked to the Marxist position on the nature of ideology this problem might have been avoided. However, the raison d'être for the work would also have been obviated. Other than a brief reference to *The German Ideology* and a vulgar reading of what the author styles the "interest" theories of ideology, no such examination is present. As in the treatment of Weber, this neglect is unfortunate, since Carlton structures his argument around an implicit and explicit dismissal of the adequacy of this tradition for the understanding of how or why an ideology is accepted. Before engaging in such an elimination, Carlton might best have consulted the work of Georg Lukács in *History and Class Consciousness* and the many who followed his work. Even the Frankfurt School and the French Situationists would have proven instructive on these matters.

The thrust of Carlton's complaint is that class-based analyses of ideology provide no explanation of the means by which an ideology is formu-

lated or how it gains hegemony (p. 25). Perhaps the ultimate defect in his work is that he does not answer his own central questions. He provides a gloss on a vast body of sociological theory and a seemingly adequate piece of social history. As analytical sociology, however, this work is a failure both in conception and in execution. It can be recommended only as social history or as an example of functionalist theory in action. Given the development of American and continental sociologies over the past decade, in the latter guise this British product is something of a curiosity.

Catholic Schools in a Declining Church. By Andrew M. Greeley, William C. McCready, and Kathleen McCourt. Kansas City: Sheed & Ward, 1976. Pp. ix+484. $15.00.

James R. Wood
Indiana University

Catholic Schools in a Declining Church focuses on value-oriented education in the context of rapid social change. Its empirical base is a replication of the 1963 NORC study of Roman Catholic schools in the United States (A. M. Greeley and P. H. Rossi, *The Education of Catholic Americans,* Chicago: Aldine Press, 1966), and it is worth the attention of sociologists in a number of specialties including educational sociology, sociology of religion, and sociology of organizations.

Chapter 1 ("Catholic Schools and Social Change") summarizes the findings of the 1963 study and some follow-ups, details the dramatic changes affecting the Catholic church during the 10 years between the two studies, and raises questions about the effects of those changes on the relationships found in 1963.

Comparison of the Catholic population in 1963 with that in 1974 (chap. 2) reveals that the new liturgy has proved quite popular; loyalty to the church is still strong; parochial schools are still widely endorsed; and, while church attendance is down among adults, communion reception is up. However, the priesthood is viewed less positively and acceptance of the Catholic sexual ethic has plummeted.

Chapter 3 ("Demography of American Catholicism") uses composite data from 12 national surveys to describe American Catholics, concluding that "there is no such thing as a 'Catholic factor' that impedes educational, economic, and occupational success" (p. 69). Catholics' lower scores on indicators of success resulted largely from the immigration experience. In their achievement Catholics under forty now trail only Presbyterians, Episcopalians, and Jews.

While Catholics' religious beliefs and practices are changing, traditional Catholic political ties remain strong. Chapter 4 ("Political Continuity and Change," by Joan L. Fee) demonstrates that Catholic loyalty to the Democratic party has not diminished with increasing socioeconomic status or

suburbanization. "Council or Encyclical?" (chap. 5) seeks an explanation for the changes in religious values in the Catholic church, attempting to establish links between historical events (and changing attitudes associated with them) and the decline in religious devotion among Catholics. The principal conclusion is that the Second Vatican Council's positive influence on Catholic belief and practice since 1963 has been overshadowed by the decline in acceptance of papal leadership and of the church's sexual ethic. The authors argue convincingly that most of the decline in Catholic religious belief and practice since 1963 can be traced to Pope Paul VI's tradition reaffirming birth control encyclical—*Humanae Vitae* (issued in 1967).

Catholics who attended parochial schools are more likely to accept changes in the Catholic church (chap. 6). However an earlier relationship between sexual orthodoxy and parochial school attendance has completely disappeared. The percentage of all school-aged children from Catholic families attending Catholic schools has dropped from 44% in 1964 to 29% in 1974 (chap. 7). It is argued that the decline does not reflect a lack of support among Catholics for such schools but the fact that the schools have become either financially or geographically inaccessible to them. In fact 80% of the respondents reported a willingness to increase their contributions to the church in order to keep Catholic schools open (chap. 8). Assuming these increases are there for the asking, the authors calculate that about 1.8 billion dollars are untapped by the hierarchy because of its unwillingness to adapt the schools to the contemporary environment.

Catholic-educated Catholics are more likely to have favorable attitudes toward the Catholic clergy (chap. 9), and parochial schools have had some success in influencing the values of those who attended them (chap. 10). Persons attending such schools, for example, are more likely to express the Christian value of "hopefulness."

The book's general conclusion is that Catholic schools can play a crucial role in stemming the decline of the church and preparing Catholics to live in a changing society. The church leadership should be strengthening its schools; it is doing exactly the opposite.

The book puts to rest the notion that there is a Catholic factor holding American Catholics back in their economic achievement—a major contribution. However, the important theoretical issue is still unresolved. Is there simply no relationship between the Catholic ethic and economic behavior, or have Catholics taken over from American society an achievement ethic that has supplanted the previous Catholic orientation discussed by Max Weber and Gerhard Lenski? Focusing on the comparison groups and assuming that religious ethics do influence economic behavior, another angle to be explored is whether there was a decline of the Protestant ethic in a generation reacting to the philosophy of parents who had "made it" in an affluent society.

While the book does not discuss organizational theory it provides the materials for reflection on an important organization (the Catholic church) as an open system permeable by the environment. The book provides an

excellent example of organizational managers failing to adapt appropriately to the environment because of a flawed communication structure.

There are a number of annoyances in the book. The absence of an index in a 493-page book is indefensible, all the more so since the writing is loose and often repetitious. Also, although there are more than 150 tables and figures in the body of the book alone, there is no list of tables and figures.

The sometimes tedious explanations of technicalities not only impede the social scientist's reading but also are unlikely to help most general readers. Both would have been better served by a more succinct text followed by technical appendixes. On the other hand the appendix projecting the future would be of special interest to the reader who is least likely to read the appendixes. Nonetheless with more than 300 pages of text and 50 pages of appendixes (including the complete questionnaires) few readers will feel cheated.

Finally, some readers may object to the polemic character of the book (it is not confined to Greeley's personal "Afterword"). However, the data are solid and the evidence for most of the interpretations is convincing. Greeley et al. should not be faulted for treating sociology as though it mattered.

The Wayward Gate: Science and the Supernatural. By Philip Slater. Boston: Beacon Press, 1977. Pp. xviii+238. $9.95.

Michael S. Kimmel
University of California, Berkeley

Philip Slater has consistently been a sociologist slightly ahead of his time. Earlier books developed reasoned and readable critiques of contemporary American culture, human sexual mores, and the burgeoning of the technocratic ethos. *The Pursuit of Loneliness* (Boston: Beacon Press, 1976), which became a best seller, established him as a major cultural critic.

But as much as he has been ahead of the times, Slater has also tended to remain on the fringes of academic legitimacy. His keen sociological eye has focused on trendy topics, and his clever turns of phrase have been cast as "pop-sociology"—the kiss of death to readers of academic journals. Finally, Slater has disassociated himself from the university, preferring to devote himself to full-time writing.

In *The Wayward Gate,* Slater turns his attention to the supernatural in a sustained polemic against the "armchair skeptics" of the scientific establishment. Rather than ahead of his time, Slater is here jumping on a bandwagon to catalog the cosmic events science seemingly cannot explain. His central question is, "What is it about some events that allows them to be accepted, while others—reported by the same caliber of witness—are rejected?" (p. 52).

But *The Wayward Gate* is not a systematic study of any set of supernatural phenomena; rather, it is a veritable smorgasbord of events that defy rational explanation. We are invited to nibble at short descriptions of UFOs, sea serpents, ESP, mental telepathy, yetis, Atlantis, ghosts, out-of-body experiences, and reincarnation without digesting the evidence concerning any one. Like us, Slater is still a voyeur, observing all these phenomena from the outside.

A study of supernatural phenomena is not Slater's goal. Instead he is attempting to lay the groundwork of a critique of the assumptions that make the possibility of such occurrences unacceptable. While scientists examine supernatural evidence with rational skepticism, Slater approaches each event with wide-eyed innocence that assumes "Why not?"

Slater's interest, then, is a critique of the way the scientific community deals with the supernatural. He seeks to tear the veneer of objectivity from scientific study, to show that the "Emperor of Rationality" stands naked in the face of evidence of the supernatural.

But even this critique of the scientific establishment is not the end point of the book but another route that allows Slater to develop his most central concern. "Science," he writes, "is founded on the suppression of feeling, on the denial of motivation, and on the need to control" (p. 115). It is biased, he argues, "against anything that threatens the domination of rational control and the power of the human ego" (pp. 95–96).

The Wayward Gate is a morality tale, and the villain is the rational ego. The ego promotes a "Tinkertoy mentality" in which "experience is most safely defined as a collection of little separate particles which can be brought together into larger particles and taken apart again. The need it expresses is one of control and manipulation" (p. 145).

The scientific establishment, according to Slater, is governed by a community of Tinkertoy egos; has developed a complex set of mechanisms for handling evidence it cannot integrate including "amnesia," discrediting the witnesses, mythification, declaring it a hoax, weighing the evidence, and theoretical disproof. In spite of this, some unexplainable phenomena find their way to public acceptance. Failing at disproof, science attempts to take control of the problem, redefining earlier blindness as technical error. (Slater mentions gorillas, meteorites, and giant squid as phenomena once denied scientific legitimacy.) When science does gain control of the problem, Slater claims, it is rationalized to a boring obscurity.

These assumptions of rationality and control are far more pervasive than the scientific community itself, Slater argues. They are the defining cultural characteristics of Western society. The rational ego, he asserts, has reorganized Western culture; we are now ruled by an "Ego-Mafia," a "collection of dictatorial egos, huddling together in a smoke-filled room, trying to extend, through concerted action, their stranglehold over their respective territories" (p. 215).

What Slater has written, then, is a three-staged critique of American culture. He has first used supernatural phenomena to criticize the scientific establishment. This critique became the means to uncover the assumptions

of rationality, objectivity, and control. The critique of these assumptions is finally a critique of contemporary American culture.

Slater has developed these critiques in an interesting and intelligible way. Unlike so much recent sociology, *The Wayward Gate* is well written, punctuated by metaphor, and sustained by a lively prose style sensitive to the nuances of language. It is unfortunate that Slater has interspersed a series of "fables"—short morality tales that are intended to persuade by allegory. While they are interesting little vignettes, they are unconnected explicitly to the text and contain somewhat obscure messages.

In the end, Slater goes further out on a limb than the substance of his cultural critique might allow. He predicts a virtual "Age of Aquarius," a renaissance of egolessness, which will dispel our destructive conceptions about the centrality of matter, the linearity of time, and the imperious control of evolutionary progress.

In such predictions, Slater is one of a rare breed. In a decade when so many sociologists are restricting their vision to become indistinguishable from other new professionals, he remains one of the last sociologists of the sixties. And even if he does overstep his critique, it is reassuring that a sociologist of Slater's caliber is still willing to take intellectual risks.

World Conqueror and World Renouncer: A Study of Buddhism and Polity in Thailand against a Historical Background. By S. J. Tambiah. New York: Cambridge University Press, 1976. Pp. vii+557. $37.50.

Roger E. Harmon
University of Washington

S. J. Tambiah's *World Conqueror and World Renouncer* is a major contribution to the study of a number of complex issues. No scholar to date has attempted either as comprehensive an account or offered as provocative an interpretation of the relationship between Buddhist religion and polity in Thailand. While the focus is on Thailand, comparisons are made between Thai patterns through time and patterns in Burma, Sri Lanka, and early India. The book will prove stimulating not only to scholars interested in South and Southeast Asian societies and religions but also to those concerned with other historic traditions and with general sociological theory.

Tambiah states that he originally did not expect to draw the extensive historical backdrop which forms a major portion of the study. However, as he began an analysis of field data collected in Thailand in the late 1960s and early 1970s, he found that he could not treat contemporary issues meaningfully unless he referred to the developments and structures of the 19th century. This was a period in Thailand characterized by reform and schism in the Buddhist Sangha ("Order of Monks"), by government restructuring under such kings as Mongkut (1851–68) and Chulalongkorn (1868–1910), and by intense Western impact. In turn, Tambiah found

that the 19th century was fully understandable only in terms of the preceding Ayutthaya Period (1450–1767). After he had worked his way back through several additional Thai historical periods, and through the critical linkages between Thai and other Southeast Asian Buddhist societies in the early centuries of the Christian era, Tambiah's inquiry finally took him to early Buddhism in India (6th century B.C.) and to a comparison of the fundamental Buddhist and Brahmanical conceptions of world and societal order. Part one, the first 12 of 20 chapters, is a detailed account of the findings of this long journey and sets forth the major theoretical position developed in the book.

The approach Tambiah has taken allows him to highlight continuities which have characterized Southeast Asian Buddhist societies through time and to relate transformations to these continuities. For Tambiah, the changes which have taken place are best viewed not as linear but with reference to the "recurrence of structures" (p. 5). Throughout the book we are shown how numerous changes are not actually the substitution of one system for another but are variants of basic, long-existent patterns—patterns which are a possibility of the system itself. Central to this argument is Tambiah's thesis that from its very beginning Buddhism included a conception of proper political action and structure, with "kingship as the ordinating principle of the polity cum society" (p. 515). This view contrasts with the assumption that Buddhism was originally a *religious* doctrine for salvation seekers, and only in the time of the Indian king Asoka (3d century B.C.) did there develop an association of Buddhism with a particular political structure. Some will conclude that Tambiah has underplayed change in his effort to stress continuities and will reject his major thesis. Regardless, his treatment of the tensions *within* religion, *within* polity, and the tensions and linkages *between* religion and polity at various periods of Thai history constitutes a major contribution in itself. Others may be more interested in the models of Thai society he develops—the galactic (concentric circles) and radial (center dominated) polities. Some readers will want to focus their attention on part two of the book, which deals with contemporary Thai Buddhism and society, especially the Thai Sangha and its relations with the state.

In part two, Tambiah extensively sets out his findings from recent fieldwork on urban Buddhism. His account, in combination with several recent, related studies of urban Buddhism and with the wealth of village data available, will allow a much greater understanding of Thai Buddhism and society than has been possible to date. Whether or not Tambiah's book itself represents an acceptable synthesis is already causing a great amount of debate among scholars specializing in Southeast Asian studies. Tambiah includes large amounts of data, but his selection is uneven. At some points there is an over reliance on lengthy passages from Western scholarship of the 1930s and 1940s, where what is needed to build a strong case is original research using Thai sources and greater attention to more recent scholarship. Furthermore, at various points in part two, one is left with the feeling that field observation of some of the phenomena discussed is

needed. For example, following discussion based largely on documentary evidence, interview data, and observation in Bangkok regarding the role of the monk in Buddhist propagation and development activities, Tambiah writes: "But what looks like feverish excitement and concentrated effort in Bangkok may well look like dissipated energy making a weak impact in the vast arena composed of countless dispersed villages" (p. 446). It may indeed, and an assessment based on observation of the monk's participation in these specific activities would seem required to support the model developed regarding the relationship between the government and the Sangha. This points to one of the major limitations of the book, namely, at several junctures we are presented with an account of what the national government and/or Sangha leadership states that they wish to achieve plus a sketch of the institutional apparatus used, but significantly less on what is, in fact, occurring at the village, district, and provincial levels. In previous research Tambiah has proven himself to be extremely skilled at participant observation; his talents in this regard could have been put to more extensive use in collecting the data for this book. Finally, some will be surprised that a book of this nature has but two bibliographical references to works in the Thai language and that it contains numerous incorrect and inconsistent transliterations of Thai words. These will be reasons why the book will not gain the immediate acceptance by numerous scholars working on Southeast Asia who have taken more limited research foci and produced studies of less breadth. Yet, it must be said that Tambiah has broken new ground in a number of ways—especially in his effort to combine textual and contextual approaches. I hope that the problems mentioned above and the book's length and cost will not keep it from the hands of scholars. Many will appreciate this major attempt at synthesis and will find much of theoretical and methodological interest.

Religion and Conflict in South Asia. Edited by Bardwell L. Smith. Leiden: E. J. Brill, 1976. Pp. 117. Fl 42.

Trevor Ling
University of Manchester

Social conflict in South Asia is often identified as religious conflict. In India, for example, tension and open conflict occur not only between Hindus and Muslims but also between caste Hindus and Untouchables or tribal people. A similar situation exists in Bangladesh, where the Hindu minority have noted without much confidence that the government declares the state to be secular, that is, religiously noncommitted. How far these tensions which undoubtedly exist between social groups in South Asia are adequately or even properly understood as religious in character would be clearer if there were agreement about the meaning of "religion." Simplistic one-dimensional analyses in terms of Hindus versus Muslims do not, in any serious way, contribute to our understanding of social reality in

Asian countries. Fairly early in modern historical studies of Bengal it was demonstrated that the socioeconomic locations of the contending parties was a factor of at least as much heuristic importance as their religious identities. It could matter less, for purposes of accurate analysis, that Muslims were engaged in violent action against Hindus than that Muslim *peasants* were engaged in such action against Hindu *landlords*.

A volume whose title connects religion and social conflict arouses expectations of this kind of elementary refinement of social analysis. These are only partially fulfilled by the collection under review. The two papers that address themselves most directly to the relation between the religious and the socioeconomic dimensions are A. J. Salvadurai's "Land, Personhood, and Sorcery in a Sinhalese Village," and H. L. Seneviratne's "Aristocrats and Rituals in Contemporary Ceylon." Salvadurai's paper is, he says, an attempt to pave the way for an alternative conclusion to that drawn by Edmund Leach in his study of Paul Eliya that "complicated arguments about . . . kinship and status were a way of talking about a quarrel over land and water." To Salvadurai, this conclusion implies the basically decisive nature of material realities. He argues, therefore, from two lengthy disputes about land in the village of Mulgama in favor of the opposite conclusion, namely, that these are really disputes not about land but about "personhood," and in particular between "whole persons" (indigenous villagers) and "part persons" (outsiders). On the evidence presented, his argument looks strongly like the kind of exercise Max Weber eschewed, that is, substituting "for a one-sided materialistic an equally one-sided spiritualistic causal interpretation of culture and history." Seneviratne's paper (only four pages long) is an account of the use made by Kandyan aristocrats of certain traditional religious rituals centering on the Temple of the Sacred Tooth to enhance their own political and social status. Brief though it is, this paper presents the most directly social analytical material in the volume.

One other contribution with a somewhat similar perspective is Cyriac K. Pullapilly's study of the Izhavas of Kerala and the attempt made by this community of Untouchables to improve their status in relation to caste Hindus. Pullapilly's approach is primarily historical—the first half of the paper presents a well-documented social analysis of the Izhava community's history, and the second half summarizes the action the community has taken to improve its social condition since about the end of the last century. Based to some extent on interviews and field work, it will rightly find a place in the literature of social change in low ranking communities alongside such work as Bernard Cohn's on the Camārs of North India.

B. G. Gokhale's contribution deals with the movement against untouchability led by Dr. Ambedkar and places this movement in its sociohistorical context in Maharashtra. Gokhale shows how Ambedkar saw untouchability as an essential correlate of caste and attacked it directly after he returned from England and the United States as a political leader of the Mahar community. His intention from the start was to lead the Untouch-

ables out of Hinduism altogether, for, unlike Gandhi with whom he had strong disagreements, he did not believe the caste-system could be reformed. It could only either remain as it had always been or be abolished. His attack on the system was mounted, therefore, not in religious terms but in terms of social criticism.

The Bangladesh situation is dealt with by Joseph O'Connell in terms of the dilemma which exists in a nation that is predominantly Muslim when the state is constitutionally secular. Whatever viability this arrangement possessed in 1971 immediately following independence from Pakistan during the period of bitter "anti-Pakistanism" has been eroded in the years since. With the assassination of Sheikh Mujibur Rahman (which occurred just after O'Connell's paper was written) the increasingly Islamic character of Bangladesh, noted by O'Connell, has become very evident. The social and religious tensions in Bangladesh appear to be too strong for a genuinely democratic and secular state to remain stable. Yet in the swing toward an Islamic type of state it is further *social* tension which will inevitably be created.

There is an excellent account by Robert D. Baird of the difficulties faced by constitutional lawyers in India as they attempt to apply the principles of the secular state to a situation in which socioeconomic stresses distort and skew conventional religious boundaries and make the separation of the religious and secular virtually impossible. Finally a useful piece of unstructured reporting of the 1975 insurrection in Ceylon by Agehananda Bharati is also included.

The unevenness in length and approach of the papers in this collection is offset by Bardwell Smith's carefully reasoned introduction. It could be read equally well at the end or at the beginning, and certainly nothing is lost by reading it twice.

Caste, Class and Democracy: Changes in a Stratification System. By Vijai P. Singh. Cambridge, Mass.: Schenkman Publishing Co., 1976. Pp. viii+158. $12.50 (cloth); $5.95 (paper).

Anson D. Shupe, Jr.
University of Texas at Arlington

Caste, Class and Democracy is an important contribution to the sociology of South Asia, applying an intriguing (though sometimes frustratingly uneven) blend of anthropological and sociological strategies to the study of social mobility in three rural communities of modern India over the period 1930–65. In particular, Vijai P. Singh's analysis is sensitive to the complexities and inconsistencies of caste stratification, both as they existed before independence and as they are now evolving. His treatment of this caste dimension, integrating it with conventional Western status indicators such as education and occupation, is faithful to his subject's cultural con-

text but does not prohibit comparability with Western data. This, in my opinion, is the book's most significant accomplishment.

There are only five chapters in this slim volume, the first of which can be by-passed by readers familiar with Marxian, Weberian, and functionalist approaches to stratification. This chapter is clearly intended as an obligatory review of major classical perspectives, but in only 14 substantive pages it can hardly do more than outline general points.

The second chapter offers a review of the literature on caste stratification. In a refreshing departure from the older (but not yet totally discredited) static view that India's caste system has been characterized by rigidity and an absence of mobility, Singh brings the reader up-to-date on the dynamics of caste and individual mobility. He suggests that caste be regarded as a local phenomenon (in the absence of any national or, often, even any regional consensus as to the rank of specific *jati*) that enters with other stratification dimensions into the equation that determines a family's overall status. His summaries of Sanskritization, Westernization, and dominant caste concepts are succinct enough to qualify this study as an appropriate supplementary text for a graduate stratification course.

As a qualification to this otherwise enthusiastic endorsement of his study, however, there are several methodological omissions, particularly in chapter two. For example, an appendix presenting the interview schedule (and sociolinguistic details thereof) used in the survey portion of the study would have been helpful to the reader. And Singh is ambiguous in regard to such a basic datum as the total *N* of his survey. We are told the size of the village populations and that all heads of households during his 1930–65 fieldwork were designated as respondents (hence attempts were made to interview them), but return rates and similar information are not reported. A similar omission occurs when we obtain a rough but imprecise idea of the number of local "judges" (informants) who aided him in enumerating and ranking *jati*. In general, a more detailed description of his problems and decisions while conducting survey analysis among India's rural population would probably have provided a valuable anecdotal increment to the literature on comparative social science methodology. As others who have done foreign area field work might testify, the blind alleys and pitfalls can make extremely informative reading.

I characterized Singh's blend of anthropological and sociological strategies as "frustratingly uneven." This unevenness is largely forced on his analysis by data limitations. At the end of the second chapter the reader learns that the data characterizing the period 1930–46 "are largely impressionistic" (p. 47), that is, qualitative and retrospective, while the 1947–65 data are of more uniform quantitative quality. Thus, of his five primary stratification variables (caste, political power, property ownership, educational achievement, and occupational status), Singh admits that caste rank and political/leadership power are actually incomparable (at least directly) over the two time periods because of changes in caste ranking systems and local political structures. Property ownership is assumed to have remained relatively constant, rendering it a not very useful gauge

of inequality. Only the conventionally Western measures of educational achievement and occupational prestige have the longitudinal comparability he seeks.

Singh attempts to make the best of this methodological dilemma by drawing on numerous case studies, his own participation observations, interviews, and historical records. His arguments lack the sort of finality that a more strictly quantitative study for all 35 years would have provided, but only the armchair researcher, spoiled by neat decks of Western cohort survey data, will discard Singh's study because of such problems.

His conclusions point to an evolving picture of achievement orientation that began before independence, among these: that there has been a fair amount of upward (and downward) father-to-son occupational mobility; that education and diversified occupational participation (the latter rendering a serious blow to the caste system) are positively associated; and that, perhaps most surprising, an increase in the interrelatedness of stratification variables appears to have occurred during the 35-year period. In particular, he predicts that caste will survive as a social consideration but largely through the family's indirect function of social placement. Singh makes no simplistic forecasts for the Indian stratification system. Caste influence is not constant throughout India and it interacts everywhere with such dimensions as the rural-urban continuum. In a somewhat repetitive final chapter, he reminds the reader of this caveat while he speculates on caste's reduced role in an industrializing society.

Singh's study goes beyond more strictly qualitative case studies of isolated communities. Until recently, national sample data on father-to-son occupational mobility have been unavailable (and local nuances of caste prohibited it from being considered in national mobility estimates). If sociologists expect an Indian study to include comparable mobility statistics while still recognizing the influence of caste, regional analyses such as Singh's will be the best suited to provide this knowledge.

The Honorable Elders: A Cross-cultural Analysis of Aging in Japan. By Erdman Palmore. Durham, N.C.: Duke University Press, 1975. Pp. xi+148. $8.95 (cloth); $3.75 (paper).

D. Eleanor Westney
Yale University

In an effort to "broaden the base for the emerging science of gerontology" (p. 2) Erdman Palmore has produced a study of the aged in Japan—a society which, while a major industrial power, lies outside the Western cultural tradition in terms of which most theories of aging have emerged. Palmore sees Japan as a society in which the aged have traditionally enjoyed very high status (p. 2) and therefore as an ideal setting in which to examine both the independent effect of culture on social structure (p. 5)

and the theory of aging known as "disengagement theory": that people withdraw psychologically and socially as they age and that "this disengagement is *good* for both the aged and society" (p. 5). Palmore also hopes to derive some practical ideas for improving the quality of life for the aged in his own society. His book, ***The Honorable Elders***, is an important one not only for the valuable data it presents but also for the questions stimulated by his interpretations of that data.

Palmore is to be strongly commended for making this work truly comparative; all too often, comparative work on Japan contrasts Japan and the United States. Palmore, however, uses comparative data not only from the United States but also from northern European industrial societies. His empirical findings, based largely on Japanese census materials and the 1973 government survey of the aged, can be summarized as follows. (1) Health and medical care: The health of Japan's aged, as inferred both from life expectancy and self-reported health levels and Japan's health care system which provides free and readily accessible medical care for the aged, compares favorably with those in other industrial societies. (2) Income and support: With an inadequate pension system and less spent per person on older people than in any other industrial society, the aged who are self-supporting are a minority (38%) of Japanese older people (although this represents a considerable improvement from the situation in the late 1950s). (3) Family and living arrangements: Japan is unique among industrial societies in the high proportion of its older people (over 75%) living with their children. This percentage represents only a very small and slow decline over the last two decades. Only about 1% of Japan's aged (compared with 4% or more in other industrial societies) are institutionalized. (4) Work and retirement: Japan is again unique in that more than 55% of Japanese men over 65 are in the labor force. Over two-thirds of these are self-employed or in family businesses. (5) Activities and satisfactions: Japanese older people are satisfied with their present lives and anticipate rising levels of satisfaction as they age further in contrast with the decreased satisfaction levels anticipated by older Americans. The Japanese case is clearly unusual among industrial societies in terms of 3, 4, and perhaps 5. Palmore interprets these data as showing greater integration and lower disengagement of the aged in Japan, with accompanying higher satisfaction levels among them, and hence as a disconfirmation of the disengagement theory.

The data clearly indicate that Japan's aged are more involved in their children's families and the labor force than aged in other industrial societies. The major problem with Palmore's book lies in his explanation of these data. The use of culture as a residual explanatory variable has always had great appeal for social scientists looking at Japan, and Palmore has fallen victim to its spell. "The Japanese tradition of filial responsibility and respect for the aged seems to be the main force maintaining integration of the aged in their families despite the counterforce of industrialization" (p. 53), and "This [the high proportion of older Japanese in the labor force] seems to have resulted from two traditions: duty to work, and

respect for the aged which reduces discrimination in the employment of the aged" (p. 77).

Seeing culture as the causal variable creates problems for Palmore when he looks to Japan for ways to improve the quality of life of the aged in the United States. For if culture is the determinant of social structure, then in order to change the social structure one must change the culture, and this is the position to which Palmore has driven himself. "In order to improve the status and integration of Older Americans, it is necessary to improve our respect for the aged somehow" (p. 140). Instituting a "Respect for Elders Day" as a national holiday and having special celebrations for the 65th birthday are two steps he proposes.

However, Palmore's readiness to leap to culture to explain the position of Japan's aged leads him to ignore indications in his own data that "structural" factors provide a more satisfactory explanation. With well under one-quarter of Japan's older workers employed in full-time nonfamily occupations, it seems that the high labor force participation of Japan's aged is due to the persistence of the so-called double structure in Japan's economy: the continued existence of small-scale, labor-intensive farming (which engages over half of Japan's older workers); retailing; and secondary industry. Palmore also ignores the extent to which the responsibility for the care of the aged falls on the housewife, a responsibility which in Japan takes precedence over any occupational role she might find outside the home or any alternative activities she might prefer. A somewhat less rosy view of the position of Japan's aged would lead another researcher to analyze sex, income, and occupation differences in order to deepen our understanding of the effects of these variables on the attitudes and life styles of the aged.

Japan is a society which, despite a wealth of readily available data and a sophisticated social science establishment of its own, has been surprisingly little utilized as a testing ground for generating and refining hypotheses in the social sciences. It is to be hoped that the problems of a book such as Palmore's will challenge other social scientists rather than discourage them from exploring the resources of the Japanese case.

Old People, New Lives: Community Creation in a Retirement Residence. By Jennie-Keith Ross. Chicago: University of Chicago Press, 1977. Pp. xi+227. $13.50.

David G. Peck
Shippensburg State College

This book is an absorbing and gratifying work. The avowed aim of the research is to show how, in a particular residential setting in a suburb of Paris, community formation by a group of elderly persons is possible.

Jennie-Keith Ross's research focus is not on French old people but on

the possible development of a community by old people in an age-homogeneous residence. The author presents community as the key concept and quite correctly states the three major themes consistently appearing in the literature: territory, we-feeling, and social organization. These three themes are combined into a composite proposition—the higher the overall level of we-feeling and social organization, including the presence of territory, the greater the degree to which community is present. The author consistently refers her observations back to the concepts developed in the first chapter. Conceptually, the reader is never left to wander in spite of some engrossing personalities and events related in the book.

Ross begins *Old People, New Lives* by discussing both the development of community observed in diverse settings such as utopian experiments and background and emergent factors in community creation. In chapter two she describes the research setting (Les Floralies)—its evolvement, the nature of the structure, social characteristics of the residents, and her entrance into the setting.

In succeeding chapters, Ross carefully outlines the nature of social participation in the residence (i.e., work); social status and community borders (i.e., measurement of social status); the development of community norms (i.e., sexual companionship and domestic arrangements); socialization into the retirement residence (i.e., political identity); and, in perhaps the most intriguing chapter, the development, resolution, and stabilization of conflict in the residence. Chapter eight addresses the most important issue in the book: Do the characteristics of Les Floralies make it appropriate to call it a community? In terms of territory, social organization, and we-feeling the author presents convincing data to show that indeed a community has been created. By definition, residents share the territorial aspect of community. The regular patterns of social living which define a social organization are also stated to be present, for example, the development and strength of the residents' committee as an institution, and the inevitable (for most residents) political faction alignments and subsequent attitudes and behaviors. There is a significant basis for we-feeling in the residents' common historical, social, and physical age, and worries about the financial status of the residence. Ross pointedly states that the emergence of community at Les Floralies may be explained without great emphasis on the residents' age.

In the final chapter, she takes pains to provide the reader with evidence that Les Floralies is not confined to the status of a case study by comparing her research with research done in four other residential settings for the elderly, at least three of which are set in American society (the fourth's location is not easily discovered). On various factors (i.e., homogeneity, lack of alternative, investment, and irreversibility) Ross concludes that the Les Floralies experience parallels that of the creation of community in the four other settings, although the author admits that information on each factor is not available in those four studies.

This book was written by an anthropologist and is one of the still very few studies that systematically describe life inside a retirement residence

through participant observation. The research is important, useful, and much needed. The orientation of the author seems clearly to be structural functionalism, and, in most cases, the analysis seems particularly appropriate for the study of creation of community. It is not pertinent here to launch into a full critique of structural functionalism, particularly from a sociological viewpoint. Clearly, the most enlightening information from this book is that movement into a residence for the elderly does not necessarily imply fear, degradation, terminality, or whatever—even in clearly segregated housing. Given a reasonable possibility of choosing one's lifestyle, many elderly can live satisfying lives.

The book, however, does have a few problems. Ross continually states that there are significant internal and external borders for the community and that one's past has little to do with the present inside the residence. Less than a convincing point is made regarding what is the most influential variable in the analysis: political identity. Political factionalism is consistently the most influential variable in Les Floralies. One suspects that this is indeed a carry-over from the residents' past. Perhaps there was not the intimate face-to-face contact with the opposition in the past, but it appears that the elderly in French society are more politicized as a way of life than, for example, the elderly generally are in the United States. Why *should* there be less politicization in a French retirement residence?

Further, the author presents much evidence implying that the quality of life inside the residence has, at the very least, remained equal to that outside. But just how much choice do the residents have? Political factions seem to lock most residents in from the first minutes they arrive at the residence. And if a resident is not aligned with one of the factions, either immediately or eventually, that person generally remains an outsider. If one premise is that quality of life is partly a function of choice, then what *really* exists in Les Floralies?

As with most participant-observer studies, the reader struggles to feel for what is being described here. But that is not necessarily a just criticism because that is essentially the nature of the methodology. This book is generally well worth the effort of reading.

Lebanon in Strife: Student Preludes to the Civil War. By Halim Barakat. Austin: University of Texas Press, 1977. Pp. xiii+242. $12.95.

Ayad Al-Qazzaz
California State University, Sacramento

The sixties and seventies witnessed the publication of countless books and articles on student politics in this country and abroad. However none of these publications was devoted specifically to student politics in the Middle East. Thus, a full-size book dealing with this subject in Lebanon represents a healthy development and a step toward broader comparative anal-

ysis. *Lebanon in Strife* is also important for another reason. Lebanon is a country still suffering from the consequences of a civil war which tormented the country for 16 months—a war in which over 50,000 persons were killed and a larger number wounded or made homeless; over half a million, mostly middle class persons, left the country; and billions of dollars worth of material was destroyed. Although the book deals with student activism, it devotes a good part to the forces and conditions that unleashed the civil war, as the subtitle clearly indicates.

The book consists of a preface, nine chapters, and an appendix. The chapters are grouped into three parts: "A Perspective on Youth and Change in a Mosaic Society," "Sources of Radicalism and Political Alienation and the Influence of Religion and Family," and "Social Class and the Development of the Student Movement in Lebanon." Halim Barakat covers several important points, three of them worth emphasizing in this short review.

First, Barakat states that the best perspective from which to comprehend student activism is the radical approach in which student involvement in politics is viewed as a function of the established social order and its socioeconomic political structure. More specifically, student activism reflects the gap between developed and underdeveloped societies, the gap between privileged and deprived within a society, the institutionalized mechanism for maintaining the status quo, and the conditions and situational factors that allow for consciousness of and sensitivity to human suffering and injustice. The author, on the other hand, rejects in brief fashion both liberal and conservative interpretations of student activism as lacking the tools for studying activists and a sound theoretical perspective.

Second, Barakat makes an interesting theoretical distinction between a plural and mosaic society. The plural model implies harmonious relationships among several interested religious and/or ethnic groups within a unified social order. The mosaic model by contrast posits diverse groups interacting within the confines of a common structure but without consensus on fundamentals. Lebanon, according to Barakat, falls into the second model and that explains the civil war. Lebanon is characterized by lack of consensus on fundamentals, lack of extensive and open dialogue, domination of public loyalties by private loyalties and interests, geographical concentration of different religious communities, nonseparation of religion from the state, existence of conflicting reference groups, pyramidal social structure, and, finally, the absence of a unified educational system. This overall lack of consensus and unity unleashed the mass violence of the civil war.

Third, student activism in Lebanon is not new but dates back to the late 19th century. However, student interest in politics gained some momentum and became more radicalized in the past decade or so as students became more aware of the dominant system, its structure, and the cultural orientations which are essentially opposed to change. Students discovered that the liberation of their country depended on a basic transformation in their attitudes, value orientations, views, and conceptions, as well as

change in the very fabric of Arab society and its political and social systems (p. 155). However, student activism tends to be fragmented and not very effective. It suffers from many limitations. The most important one is a lack of cooperation and sharp separation along vertical loyalties (identification with religious, ethnic, regional, or tribal ties), which continue to constitute a higher significant force in Lebanon relative to horizontal loyalties, which refer to socioeconomic criteria.

The book suffers from several shortcomings, two of which I will mention here. Barakat did not use a uniform sampling procedure, did not conduct his surveys at the same time, and he did not use exactly the same questionnaires in three different universities. Such variation in technique, time, and wording is bound to affect the findings in one way or another. Furthermore, the academic background of the students in the three samples is not the same. One sample consists of students drawn from law, social science, and education while the other two samples consist of students with various academic backgrounds including science and engineering. Again, different academic orientations affect student activism differentially, as students in social science and humanities tend to be more aware of the social and cultural forces impinging upon them than are science and engineering students.

In the introduction and elsewhere the author states that his purpose is to study student politics in Lebanon as part of a comprehensive concern with development and transformation in Arab society. Throughout, however, Arab society receives a marginal and passing treatment. Furthermore, Lebanon's unique characteristics mean that, without extreme caution, it cannot be employed as a model to analyze Arab society in general.

Despite these shortcomings, *Lebanon in Strife* represents a contribution to the field of comparative politics, and thus I strongly recommend it to all students of this field and to students of Middle East society as well.

The Plug-in Drug: Television, Children and the Family. By Marie Winn. New York: Viking Press, 1977. Pp. xii+231. $8.95.

Rose K. Goldsen
Cornell University

The image you see on your television screen is produced as a cathode ray tube scatters points of light in predetermined patterns. It is similar to the principle of producing newspaper photographs by arranging black dots against a white background. The television image, however, is in continual motion. If you could photograph the picture on your screen at 1/2000 of a second, the developed print would show just a single row of bright points. Discerning the TV image, therefore, is always a collaborative effort between viewer and set. In order to pick out the likeness of Kojak or Cronkite, Mr. Whipple or Bugs Bunny, you must stop those shifting points of

light and connect them in your mind's eye where the image is reconstituted.

Television sets in this country are active about seven hours a day on the average, and American children spend hours with their eyes glued to the television screen. Infants in the cradle watch those moving points of light. Preschool children average about 30 hours a week, grade school children a few hours less. Only one activity commands more of the children's time: sleeping.

Images on the screen exert compelling fascination even for adults. For children, their power to catch and hold attention seems irresistible. Most children will watch anything and everything. Many are "zombie viewers." The term was coined by the team who developed "Sesame Street" to describe the way certain children watch—silent, immobile, mesmerized. Even the test pattern keeps some of them in the same trancelike state.

Television has been hailed as an educational tool that can provide the infrastructure children need to develop verbal skills, reading, and cognitive thinking—the kind of mental process that moves step by step, from one proposition to another. Fitting those bright dots together, however, involves mainly the brain's right hemisphere, which commands nonverbal, visual, and spatial components of thought. No matter what lessons educators may build into the television programs they label "educational," television is a nonverbal experience for the children. It teaches them neither to master their language nor to love it. It demonstrates how billboards are read, not books. Watching is the principal lesson taught by this machine which demands of the children only their passivity.

So invasive is the television experience that it has become one of the primal experiences that mold character. Adults socialized in an era before the set moved into so many homes can evaluate the television experience, casting it up against a backdrop of their own associations and real life recollections. For children raised by and with TV, experience with the machine is primal, on a par with nursing, excreting, cuddling, and infantile sexuality.

Through the ages children have played together in games that allow them to internalize the rules which keep the social system functioning. Now, television time replaces hours of playtime, and games like marbles, tops, jacks, as well as equally ancient jumping and running games, are fast disappearing from the culture of childhood. Even sandlot baseball gives way to games organized by adults who do their best to model the proceedings after professional, made-for-TV spectacles.

These are a few of the major changes that occur as the TV takes over important functions of child rearing. The implications for brain development, optical function, cognitive style, character formation, and social structure are surely revolutionary. It is what Marshall McLuhan has been telling us for years: The medium *is* the message and communications systems are extensions of nervous systems. Dorothy L. Cohen and other pioneers in early childhood education have warned that children will not

develop normally if they are robbed of play and mesmerized into passivity, learning to interact with a machine instead of with other human beings.

Now, along comes Marie Winn whose book weaves all these strands together. She leaves no doubt that the official media research establishment—in universities, on Madison Avenue, in foundations and government agencies—has spent years asking the wrong questions, pretending that television's effects can be understood by testing whether intentionally engineered "messages" hit or miss human targets who then do or do not react according to design. Such "effects," according to Marie Winn, are the least of it, overshadowed by what happens as children and machine are linked together in a single cybernetic system.

Winn is a journalist and writer, not a scholar. Her book, eminently readable, relies on original sources in some areas, secondary sources in most, the selections judiciously made. As a mother who has raised two children, Winn has had the good sense to respect her own insights and experiences as well as those reported by the parents, teachers, and specialists she interviewed for the book. I show her work to students as an inspiring example of how one person, working and studying alone and refusing to be bedazzled by conventional wisdom, can make a major breakthrough.

In her insistence on calling attention to the effects of the total television experience on brain, nervous system, and interaction patterns, Winn ignores television's massive alteration of the symbolic environment whose role in the formation of human consciousness is surely as important as the machine's stimulation of optic nerve and brain. The omission detracts little from the book's value. No book can or should try to explain everything, and this one has put together enough good ideas to keep Ph.D. candidates busy for decades.

I recommend *The Plug-in Drug* as required reading in many social science courses, especially those that overemphasize methodological hardware. Laymen should read it too, including those in government agencies charged by law to see to it that broadcasters help, not hinder, the public interest.

Literary Culture in France and the United States[1]

Priscilla P. Clark
University of Illinois at Chicago Circle

A series of cultural indicators reveals the similarity of general population literary preferences and the dissimilarity of elite literary tastes in France and the United States. The distinctive associations of literary activities with French and American elites are responsible for the distinctive characters of French and American literary cultures. Whereas French literary culture is sustained by strong ties to centrally located governmental and traditional elites, American literary culture is defined by its support from dispersed private elites. As these ties to elites define literary culture and influence literature, they are supported by the social and political organization of French and American society. Recent social and political changes have modified traditional patterns of support for literature and have attenuated but not effaced certain distinctive features of French and American literary cultures.

> The relations that exist between the social and political conditions of a people and the genius of its authors are always numerous; whoever knows the one is never completely ignorant of the other. [DE TOCQUEVILLE 1958, p. 63]

I

Associated with elites in premodern societies, the arts in modern times find themselves in a world often indifferent, and at times frankly hostile, to anything labeled "elitist." The resulting tension between the traditional elite orientation of the arts and a certain populism in contemporary mass societies is fundamental to the dynamics of art today. For although one can argue that never as at present have the arts so flourished, concerned so many, or received so much support, because they cannot depend on market support, by necessity as much as by inclination the arts seek out elites today as in the past. Perhaps more important still, beyond the crucial economic support they provide, the elite ties of the arts sustain artistic cultures—

[1] It is a pleasure to acknowledge assistance, both intellectual and material: research grants from the University of Illinois at Chicago Circle; research assistance from Michelle Fox, Elizabeth Gross, Lee Ann Nicholson, Greg Pittsley, Michael Ralston, George Rezek, Judy Wattelez, and Lucie Yassa; interviews and information from French officials in Chicago, New York, and Paris; and criticisms and comments from too many quarters to be enumerated lest I be forced to declare intellectual bankruptcy.

those values, norms, symbols, and patterns of behavior that create a distinctive milieu for the practice, pursuit, and appreciation of a given art.

I shall, however, argue the case only for literature and literary culture, and only for France and the United States. In very large measure the distinctive literary cultures of these countries derive from the very different elite associations of literature in each country. With roots in an aristocratic, monarchical, and highly centralized society, French literary culture is characterized by long and close associations with those traditional elites, notably governmental, which have been central to and especially prominent in French society; this in marked contrast to American literary culture, which evolved in a society that explicitly rejected aristocratic models of society and literature. The elites associated with American literature have been located in the private rather than the public sector and dispersed over the country rather than concentrated in the capital. The two literary cultures could not but differ profoundly as indeed they still do, despite recent modifications of established patterns of elite interaction which have attenuated though not effaced their once very great differences.

To explore these literary cultures I shall make use of what may be termed cultural indicators. Whereas economic indicators generally gauge the performance of the economy and social indicators typically point to characteristics of populations, cultural indicators aim at defining cultural attitudes, ideals, and dispositions. As social and economic indicators often coincide, so cultural indicators will frequently intersect with each of them. Thus whether given data are used to construct an economic, social, or cultural indicator will depend on what is wanted from the analysis.[2]

II

The distinctive literary tastes of French and American elites are all the more striking in view of the quite similar literary preferences of the general French and American populations. Several indicators point to this similarity: the number of titles published annually, sales of books, reading habits, status of writers, and total moneys in direct support available to writers.

Although France publishes more new titles annually (52.2 per 100,000 inhabitants in 1973) than the United States (39.7 titles), on a world scale

[2] George Gerbner and his associates have used the term quite differently in their analyses of the composition, structure, and effects of the "messages" of television dramatic programs, principally with regard to the incidence of violence. In this usage cultural indicators refer both to the "cultural" content of television programs and also to the broader sociocultural views reinforced by this content (see Gerbner 1973 and Gerbner et al. 1978). "Cultural indicators," as I use the term, are concerned neither with media content nor with its effects but are used to define the institutions, behavior, values, and symbols of a particular milieu or (sub)culture, in this case the ambient culture of literary activities.

both countries are in the same upper-middle range.[3] And if, of all books pu lished, France continues to publish a higher percentage of literary titl again on a world scale the United States does not lag that far behind nor France by any means the leader. The proportion of literary titles has d clined in both countries: from 45% of American publications in 1811–2 literature decreased to 42% of all titles in 1900, 35.9% in 1937–39, 37% i 1950, 35.9% in 1960, and 22.3% in 1973; France published 28.3% literar titles in 1938, 31.5% in 1950, 35% in 1960, and 26.2% in 1973.[4] As i other industrialized countries the increase in scientific and technical title has counterbalanced the relative decrease of literary works.

In neither country are books a major commodity: in France and in th United States, sales of books represent a tiny and virtually identical propor tion of the GNP (0.28% in 1970 for the U.S., 0.26% for France).[5] And since the French on the whole seem to read about as much (or as little) as Americans,[6] we should probably not be startled at the virtually identical scores obtained by writers in France and the United States on the NORC Occupational Prestige scale.[7] Finally, relative to their populations, France

[3] Based on 1973 totals from the *United Nations Statistical Yearbook* (United Nations 1976) (French totals include pamphlets). Titles per 100,000 inhabitants ranged from 0 9 to over 100. Over half (51) of the 81 reporting countries publish fewer than 20 titles per 100,000 inhabitants This and the other cultural indicators are discussed in greater detail in Clark (1979).

[4] The 19th-century average is reported in Charvat (1968, p. 35); the 1900 figure is from Hokkanen (1969, p 67); later American and all French figures are taken from the *United Nations Statistical Yearbook* (United Nations 1954, 1963, 1973, 1976). The United Nations "Literature" category (number 8, Universal Decimal System) partially covers the following categories of the American Publishers Association· "Fiction," "Juvenile," "Literature," "Poetry, drama " Other designations of "literature" or "literary" (prizes, book reviews, television programs, names of streets) represent my own judgment.

[5] Sales in the United States in 1970 were $2,922,000,000 or 0 28% of the GNP Sales in France in 1970 were Fr 1,935,000,000 before taxes, or $450,000 000 ($1 00 = Fr 4.30), 0 24% of the GNP; Fr 2,069,000,000 or $481,160,000 with value-added tax included, 0 26% of the GNP.

[6] Ennis (1972, p. 51) reviews reader surveys of adult Americans (1949–62), which showed a range of 46%–60% of those surveyed having read a book within the previous year. A 1973 survey in France showed that 69 7% of the sample had read a book within the previous year, but the survey included those in the 15–19 age group, who read more than adults, whereas the American samples were limited to adults In the most thorough comparative study of reading habits, the French, who averaged seven minutes per day reading books, did not differ greatly from the Americans, the sample of whom spent five minutes on books Of the French sample 12% read books compared with 7% of the Americans. The differences exist, to be sure, but they are not great, especially in comparison with the 29 minutes a day the Russians spent reading books or the 17 minutes of the runner-up Poles (Szalai 1972, p. 580).

[7] On a scale of 100, "auteur de roman d'avant-garde" ranked 55 7 in France (Clark and Clark 1974) as compared with 57 0 for "author" in the United States, 64.6 in Belgium, 66 in Poland, 70.6 in Israel, and 71.0 in Italy (Treiman 1977). Whatever the arguments against the Occupational Prestige Scale, they presumably apply in one country as another.

and the United States in 1975 provided about the same amount of direct support for writers in fellowships, prizes, and grants (Clark 1976, 1977).

Should we expect otherwise? Both highly industrialized and technologically advanced, France and the United States are also similarly characterized by all but universal literacy, more or less market economies (publishing is not a nationalized industry in France), and more or less representative governments based on universal suffrage. Far greater contrast should be expected between either France or the United States and most so-called Third World or developing countries which are neither technologically advanced nor universally literate nor have representative governments.

The shared literary preferences of the French and American general populations would not provoke much comment were it not for a stereotype which assigns to literature, and to writers, a special and especially important role and/or status in French society—in contrast to the supposed lack of interest and even dislike on the part of the Americans whom Mencken dubbed the "booboisie." Gertrude Stein, for one, was impressed by the deference writers received in France:

> But really what they do do is to respect art and letters, if you are a writer you have privileges, if you are a painter you have privileges and it is pleasant having those privileges. I always remember coming in from the country where I usually kept my car and the garage was full more than full, it was the moment of the automobile salon, but said I what can I do, well said the man in charge I'll see and then he came back and said in a low voice, there is a corner and in this corner I have put the car of Monsieur the academician and next to it I will put yours the others can stay outside and it is quite true even in a garage an academician and a woman of letters takes precedence even of millionaires or politicians, they do, it is quite incredible but they do, the police treat artists and writers respectfully too . . . . [Stein (1940) 1970, p. 21]

One might simply dismiss Stein as out of date and out of step with indicators from the end of the 20th century were it not for more recent statements to similar effect[8] (not counting the number of persons, sociologists included, who have said to me that "of course" the French must read more than Americans, subsidize literature more, think more highly of writers, or whatever). But the contradiction is only apparent. As one might suspect, insofar as in France more attention is accorded things literary or exceptional

[8] Similar sentiments are expressed by nonliterary figures, French no less than foreign. A sampling: a German academic—"France is the only country which has a 'religion' of literature" (Curtius 1932, p. 204); an English journalist—"Often [writers and intellectuals] have been highly vocal rather than directly influential; but always [French] society has treated them as an honored minority, like the holy men of India" (Ardagh 1968, p. 354); an American journalist—It is "a popular belief that the average Frenchman is a cultural *maven*" (Lottman 1975); and most recently a French writer—"It is certain that the status of writer is especially honorable and honored in this country" (Aron 1975, p. 140; cf. pp. 134, 145, 204). Also cf. Wilson (1971, p. 156).

prestige granted literature, such honor and glory derive not from French society as a whole but, rather, from the attitudes and behavior of elites in that society.

And as some indicators point to the congruence of the general French and American populations in regard to literary matters, so others reveal the still considerable differences in the literary preferences of French and American elites. Whereas traditional elites, the monarchy, the aristocracy, the Church and subsequent governments have especially influenced literary activities in France, American literary life owes more to the untraditional or "modern" elites of commerce and industry.

III

With roots in the princely courts of the Middle Ages and the Renaissance, the elite ties of French literary culture were well established before the 17th century, but it was then that Richelieu and subsequently Louis XIV centralized patronage of literature (as of the arts generally) and associated it with a national government. If all the arts did service, literature was privileged. More clearly than the other arts, it instilled the values and norms of an elite milieu and integrated the individual into that milieu (see Aron 1975, pp. 134, 204). French was a common and exclusive language which not only favored social interaction between *littérateurs* and traditional elites but also affirmed a common elite identity (which is not to say equality). Court and salon alike stressed participation in cultural activities: courtiers and salon habitués not only listened or read or watched, they wrote. Louis XIV himself acted in plays, danced in ballets, and even turned an occasional (poor) sonnet. The genres most in vogue were typically cultivated in salon conviviality—maxims, portraits, epigrams, poetry meant to be read aloud, and plays given private readings. Not withdrawn into the solitary garret of the Romantics, the 17th-century writer belonged in a resolutely social setting in which writing did not define a separate status so much as a role.

While governmental patronage of literature did not disappear altogether after 1789, it diminished drastically (see Clark and Clark 1977; Lough 1978). No longer the usual means of support that it had been during the ancien régime, patronage of individuals became an anachronism, even an anomaly, in the 19th century. One is not really surprised that even the quite fastuous Second Empire should in 1862 budget considerably more money for prizes to horses than for prizes to writers and scholars and allocate more funds to buy horses for the imperial stables than to purchase books for the imperial library. Today the moneys disbursed in direct support to writers are windfalls for a few individuals but have little impact on literary life.

Nevertheless, insignificant as they may be in strictly monetary terms, government subsidies attest to concern for and (presumably) interest in

literature and writers. It is therefore noteworthy that the French government accounts for almost 70% of all direct support to writers in France, while the American federal government supplies only 20% of the direct subventions available in the United States (in 1975; see Clark 1976, 1977). This greater involvement of the national government in France may be largely symbolic, but it is just such symbolic concern that is so striking a feature of French literary culture.

The Académie Française is undoubtedly the best-known because longest-lived embodiment of governmental involvement in the French republic of letters and of the interpenetration of polite and literary society. Founded in 1635, the Académie was charged with fashioning the French language into the proper instrument for French society, which in the 17th century meant the king, the court, and dependent Parisian milieus. It has always drawn its members from many elites (men of letters were and still are a minority) precisely because of this injunction to model the written language on the cultivated speech of salon and court instead of the technical jargon of law or theology (Renan 1860, p. 341). In the same spirit of linguistic punctilio, the Haut Comité de la Langue Française was established in 1967 by Georges Pompidou, then the prime minister, to police the use and abuse of French on national television and radio, while the Académie Française continues periodically to censure incorrect expressions, to purify French of "le franglais" and other contemporary "barbarisms," and to devise authentic French terms as technological and other innovations require(see Bengtsson 1968). This characteristically French "language fetichism" (Bourdieu 1975) or "linguistic imperialism" (Aron 1975, p. 22) recently took Academicians from the tranquillity of the Dictionary Commission to the corridors of the National Assembly where intensive lobbying brought passage of a law which makes it a legal offense and subject to fine to employ foreign or un-French expressions "fabriqués contre le génie de la langue" (as the Académie puts it) in official communiqués: no more "hardware," no more "cockpit," no more "week-end," at least in governmental prose.

The French government is also tied to literary life through a number of literary prizes. The ORTF (Office de la Radio et Télévision Française) awards a literary prize, the Ministry of Education gives two, and not long ago, the Gendarmerie Nationale (National Guard) funded a prize for a novel that would show gendarmes—not ordinary policemen—in action. The Académie Française, the largest single source of prizes, combines private legacies and governmental subsidies to support 143 prizes ($65,000 in 1974), while the Centre National des Lettres recently established seven prizes ($19,000 in 1975).

Striking as they are, French government support and surveillance of letters assume a broader significance as expressions of a more general elite interest in literature. While governmental affiliated groups awarding literary

prizes disbursed 67% of all prize moneys (81% when the Académie Française is included), they represented only 5% of all prize associations (14 of 267) and accounted for not quite 20% of all prizes awarded (160 of 816). With a mean value of $236, literary prizes do not support individuals so much as conceptions of literature, more specifically traditional elite definitions of literature. Although the oldest (and still awarded) literary prize in France dates from 1323 and even though various academies awarded prizes every now and again under the ancien régime, prizes and the associations that awarded them did not begin to multiply until the 19th century—a period when new schools of literature demanding recognition and new publics clamoring for different literary fare were challenging time-honored conceptions of literature. It was then that mass culture first reared its (to most prize donors) ugly head. The bequests received by the Académie Française (the most notorious was for a prize to reward "useful" works) and many of the other prizes funded over the 19th century attest to a conscious defense of traditional values.

In the face of great social as well as literary change, literary prizes offered a guarantee of quality, or more accurately, of a certain, usually traditional, conception of quality. Which is why such a hue and cry is raised whenever the winner of a major prize does not live up to the expectations of publishers as well as readers, for, although the great majority of prizes have no appreciable impact on the market, a few—a very few—do. The Prix Goncourt, awarded since 1902 to the "best work in prose" published during the previous year, regularly sells 200,000–400,000 copies—a regular direct market influence not even approached by any American literary prize. The five or six other "marketable" prizes sell less, but even the 20,000 or so copies of a "difficult" Médicis (established expressly for experimental fiction) are still well above usual sales (only 10% of literary works in France sell over 10,000 copies, 0.5% over 100,000 copies). Nevertheless, and notwithstanding the sales, none of these prizes concerns or is concerned with popular literature. The annual Goncourt is, in fact, denounced annually for its snob appeal, precisely because like most of the other literary prizes it activates essentially elite preferences. The Prix Goncourt simply does so with quite spectacular market effect.

Literary prizes make a specifically literary gesture; newspapers, news magazines, and television address a general public. It is therefore significant that elite interest in language and literature finds expression in general news media. The two leading French newspapers, *Le Monde* and *Le Figaro,* regularly publish columns discussing problems of language usage and reviewing grammars and dictionaries as well as more technical works on language and linguistics. The English (or American) language arouses no comparable general interest in the United States, where such matters tend to be discussed in more specifically literary or "cultural" journals.

This exposure to literature or literary "news" is confirmed by a distribution analysis of *Time* magazine and its closest French counterpart, *L'Express*, which showed that the French weekly consistently allotted more space to cultural concerns generally and to literature in particular. "Cultural" pages accounted for 24% of all pages (less advertising and letters to the editor) in *L'Express* and 16.1% in *Time,* while book and literary pages represented 12% of *L'Express* pages in contrast to 6.9% for *Time*. Since both magazines have a fairly select readership relative to the population as a whole, and since both stress the elite segment of that readership, it seems appropriate to conclude that each reflects an assessment of what will appeal to, hence attract, such an elite readership without alienating the bulk of their readers.[9]

If these media are perceptibly oriented to an elite audience, the same cannot be said of television. Even so, French television programmed more "cultural" material (8% of all hours) than television in the United States (2.6% of all hours in New York City) as well as more literature and drama (3% of Paris hours vs. 1.2% in New York) (*Le Figaro* 1976; *New York City TV Guide* 1976). Still, although the French programs are more prominent because the choice of programs is so restricted, the match is probably a draw in view of the greater total literary and cultural hours on American television.[10] In any event, unlike American television stations whose decentralized ownership orients them to advertisers, French television stations are

[9] Five issues of *Time* magazine and *L'Express* were examined for 1973–74 and five for 1977–78. From 1973–74 to 1977–78 cultural pages remained virtually constant in both magazines, but literary-book pages declined by almost one-third for *Time* (8.4% of all pages to 6.9%) and by close to one-half for *L'Express* (16 1% to 8.3%). The 1975 circulation (paid copies) of *Time* would reach 2% of the American population; the copies printed of *L'Express* would be read by 1 3% of the French population. Even though one should assume several readers per copy (*L'Express* counts on 4.5), the journals' managements stress their elite readership: 35 5% of *Time* subscribers are managers or administrators, 27.2% have master's or doctor's degrees ("Time" *National Subscribers Study* 1976), while 26.3% of *L'Express* readers had upper level management or administrative positions ("cadres supérieurs") (*L'Express* 1974). The greater resonance of books and literature for French elites (and/or the more pronounced elite character of *L'Express*) can be gauged by advertisements: in 13 issues examined for 1978 *L'Express* had over three times as many book or book-related advertisements as did *Time* (53 vs. 16). Moreover, almost half the *Time* advertisements were for *Time-Life* books or magazines and not remotely literary, whereas 40 of the *L'Express* advertisements, or 75%, were literary (mostly novels). One might note also the book advertisements in *Le Monde* and *Le Figaro,* which appear on the front pages as well as in the weekly book section, a practice not followed by American newspapers.

[10] The statistics for New York include public broadcasting. During the week of June 12–18, 1976, New York broadcast 1,750 hours total, including 45 hours (2.6%) of "cultural" programs, 22 hours (1.2%) of literary and theater programs, and 6 hours (0.3%) of literary programs alone (*New York City TV Guide* 1976) Paris television in the week of April 12–18, 1976 totaled only 189 hours, of which 15 (8%) were "cultural" programs, 5.75 (3%) literary and theater programs, and 1.83 hours (0.97%) literary programming (*Le Figaro* 1976).

administered by government appointed officials at least as mindful of what is appropriate as of what is popular. The relative prominence of "cultural" programs on French television undoubtedly reflects a judgment that these are eminently appropriate subjects. It is obvious, however, that whatever the prominence of individual programs or the publicity given to "cultural" specials on American networks, the programs themselves are at best an oasis of culture. Even in France the officials who program television broadcasting are dealing with a mass medium.

Not so the government officials who select a country's bank notes, who are relatively free to select the images they feel most appropriate for this "cultural carrier." And this indicator is striking evidence of how very appropriate a cultural carrier literature is in France. In contrast to other European countries as well as the United States, literature does spectacularly well on French bank notes: writers account for five of the seven figures on bills currently in circulation (Hugo, Voltaire, Racine, Corneille, Pascal). The individual's portrait is on both sides of the note, against a background related to incidents or places important in his life. In contrast, all American bills show political figures, the Federal Republic of Germany uses portraits from famous German paintings, and Great Britain and Italy each claim one writer (Shakespeare, Manzoni) out of five figures.[11] Moreover, the French bank notes do not identify the writers, unlike the Italian, British, and American notes and unlike the most recent French issues with a composer (Berlioz) and a (minor) painter (Maurice Quentin de la Tour). Presumably the writers were considered recognizable, although an unrepresentative personal sample was highly inconclusive in this regard. Still, it is not entirely improbable that prostitutes should pay off cops with "des Pascal" (street slang for 500-franc notes), however unlikely it is that either party to the transaction has read a word by the 17th-century writer, geometrician, and mystic.

City officials too select cultural symbols—mottos, emblems, flags, and names of streets. Paris City Councils have concurred in their city's reputation as a literary center, for 332 streets in Paris, 6.4% of all streets, bear

[11] For comparison, U.S. bills portray· $1 00—Washington, $2 00—Jefferson, $5 00—Lincoln, $10.00—Hamilton, $20 00—Jackson, $50 00—Grant, $100 00—Benjamin Franklin, $500.00—McKinley, $1,000—Cleveland, $5,000—Madison, $10,000—Salmon P. Chase. Great Britain has the queen on the front of all bills while the reverse carries: £1—Newton, £5—Duke of Wellington, £10—Florence Nightingale or the British Lion, £20—Shakespeare. Italy has: L 1,000—Verdi, L 5,000—Columbus, L 10,000—Michelangelo, L 50,000—Leonardo da Vinci, L 100,000—Manzoni French bills, unlike the others, change fairly often, primarily to discourage counterfeiters (Napoleon was on the 100-franc note a few years ago, Voltaire and Hugo are moving out of circulation) When the Banque de France decides that a new note is needed, it commissions several artists to submit proposals and drawings, among which the director of the Banque de France chooses a figure for the new issue. The only guideline seems to be to avoid individuals who would arouse strong disapproval in official circles, in other words, who would be inappropriate cultural carriers (Banque de France, personal communication).

the names of men (and a few women) of letters. In comparison only 2.1% of West Berlin streets are named after writers, 1.9% in East Berlin, 1.2% in Boston, 1% in San Francisco, 0.9% in Chicago and in London, and 0.6% in New York. (Elimination of lettered and numbered streets and inclusion of possible but doubtful names altered the results only slightly.) By the late 19th century, when these cities added and named streets, the French could boast of a long and distinguished literary heritage as the Americans could not. Thus foreign writers name 79% of Chicago's literary streets, 71% of San Francisco's, and 32% of New York's, but account for only 5% of Parisian literary streets.[12]

It is true, of course, that the 332 favored writers in Paris include a fair number of luminaries whose lights are at present dimmed if not altogether extinguished. Still, well-known writers are in sufficient number to give a literary cast to many a Parisian promenade. Other Americans of a literary bent must have reacted like Edmund Wilson who thought it "quite wonderful" to find a street in Paris named for one of his favorite authors and like him must have reflected on the very different situation at home where the efforts to rename Washington Place for Henry James, who was born there, came to absolutely naught (Wilson 1971, p. 156).

These many activities, particularly on the part of the national government (the municipal government in Paris is far less autonomous than in American cities), bespeak a certain if symbolic attention paid to literature, and this attention is by definition that of an elite. For it is not the man in the street who names the street, nor the woman paying for metro tickets who picks the cultural heroes for bank notes, nor the viewer who presents television programs. These decisions translate, directly or indirectly, tastes and preferences not of all consumers but of government elites, while most of the literary prizes, the subjects of *L'Express,* the language and literature columns of *Le Monde* and *Le Figaro,* point to the generalized elite associations of literature in France.

These elite literary preferences affect more than elites. Clearly they serve to weaken other, nonelite definitions of literature. Until recently the literary market in France seemed faintly antiquated, at least from an American perspective. Despite technical advances and a greatly expanded market, large-scale commercialization came to France relatively late, and one reason surely lies with the perception of publishers and booksellers alike that theirs was a "noble" task, hence removed from the vulgar commercialism of the

[12] These figures are based on examination of city guide books or maps. The almost exclusive use of surnames in American cities makes identification considerably more difficult than in Paris or Berlin where the given name is frequently specified. Consultation of a dictionary of Paris streets (Hillairet 1963) corrected the original estimates, increasing the number of writer streets from 282 (5.3% of an estimated 5,418 streets) to 332 (6.4% of a definite 5,218 streets). While similar research might uncover more writers in other cities as well, it would be unlikely to alter the relative positions.

marketplace. Inexpensive paperback books with a mass distribution in France followed England by 30 years, the United States by 20, and even today populistic practices are dubbed with appropriately foreign names like "le discount" and "le best-seller." The opening of a large discount bookstore in Paris (FNAC) provoked cries of outrage and dark intimations that such un-French and unliterary practices portend the demise of French literature itself. Even more "noble" than booksellers, writers are adventurers, to use André Malraux's term, engaged in a calling rather than a mere occupation (Gouillou 1975).

The entrenched elite orientation of French literary culture has affected the diffusion of books as well as their production. Public libraries, to take the most obvious example, ought to counter elitism since diffusion to a broad public is after all their *raison d'être*. It is then striking, and quite consonant with the elite associations of French literary culture, that France should be relatively poor in public libraries. With one-fourth the American population France has only one-tenth as many public libraries: 837 in 1971 versus 8,337 in the United States in 1974 (UNESCO 1973, 1975).[13] The United States has 13 libraries with more than 4 million volumes; France has only one, the Bibliothèque Nationale, which has 7 million volumes but which, unlike the U.S. public libraries (6 out of the 13), is open principally to scholars. Most of the other important libraries in Paris (Senate, National Assembly, Ministry of Foreign Affairs, the Sorbonne, and others) serve primarily individuals associated with the institution, although once again researchers may obtain permission to enter. One did not find in France the distinctive combination of *richesse oblige* and missionary zeal (and ostentation) that prompted American elites to found libraries and other cultural institutions, from museums to universities, designed for general rather than strictly private consumption.

IV

The literary preferences of French elites and the consequent interaction between literary and other elites have sustained characteristic patterns of sociability and combativity in French literary culture, and these have in turn

[13] A general encyclopedia, *Quid* (Frémy and Frémy 1976), all of whose sources I have not checked (but which is reliable on those I have been able to verify), indicates that 20% of Americans are registered in a public (municipal) library versus 4.5% of the French and 30% of the English and Danes. In France 0.75 books are loaned annually per inhabitant versus 5.4 in the United States, 10 in Great Britain, and 7 in Denmark. France spends 2.5 francs ($.58) annually per capita on libraries versus F 37 ($8.51) for Denmark, F 18 ($4.14) in the United States. Dumazedier and Hassenforder (1962) document the underdevelopment of the library system in contrast to the United States, and especially England and the Scandinavian countries. The French government has been making a decided effort to augment public library resources of which the most spectacular example is the open stack public library at the Centre G. Pompidou (Beaubourg).

reinforced certain features of French literature. The virtually unavoidable personal interaction of writers both with elites and among themselves, at court, in salons, cénacles, cafés, frequently over a convivial dinner, supper, or tea, have tended to define literary activities as public, social events as well as, at times perhaps more than, an individual enterprise (see Thibaudet 1940). It is not surprising that this sociability and the conversation that goes with it were what Mme de Staël missed most during her exile from Paris, or that, despite the disappearance of the institutions that originally brought elites together, Parisian literary life is still, or can be if the individual so chooses, very social, very *mondaine.*

The strong ties of literature to traditional elites created definite though subtle constraints to a normative taste (see Renan 1860; Arnold [1864] 1962). At the same time, the very strength of the literary norms provoked equally strong reactions, which is why literary change has so often proceeded by attack and counterattack in a typically French pattern of innovation through highly vocal, often intemperate, rupture with the past. The literary prizes that proliferated in the 19th century intensified this competitiveness to the point where a writer like Zola naturally thought in military metaphors: "Force is everything in the battle of letters." Because many literary groups were small, cohesive, and more or less expressly established to defend a particular conception of literature, they acted to constrain individuals to define their aesthetic and ideological positions and certainly contributed to the intellectualization of French literary life. The exceptional number and concentration of these literary groups and institutions presented so many obstacles for the asocial outsider, who risked neglect by that fact alone. The inescapable confrontations of rival groups, each maneuvering in a small space to impose its own definition of the literary situation, exacerbated the conflicts inherent in any literary culture. "I am attacked, therefore I still am," affirmed Zola, more bellicose than most, no doubt, yet representative of the combative intellectuality of a literary culture which savors aggressive manifestos, ringing critical prefaces, dramatic *causes célèbres,* and resounding *coups de théâtre.*

The normative constraints of French literary culture as well as its intellectuality can be related to France's notably long and continuous literary tradition (Odin 1895; Kroeber [1944] 1963, pp. 536–41).[14] At the same time, if French literary history has no really "off" periods, so to speak,

[14] This seemingly francophile assertion requires some explanation. Other European literary traditions might be as long and/or as rich as the French but their achievements tend to be clustered and not spread so evenly over time. The exceptional continuity of the French literary tradition made French writers especially suitable for Odin's study of the social factors influencing genius (1895, p. 283). Kroeber ([1944] 1963) traced patterns of creativity from the Orient to the Occident over several centuries. Following the conventional ratings of cultural "digests" like textbooks and encyclopedias (p. 23), Kroeber concludes (p. 236) that "from before 1100 on French literature manifests a continuity of performance unequaled in Europe."

neither does it boast any individual who towers over all others, a monumental figure on the order of Dante, Cervantes, Shakespeare, or Goethe. French literature revolves around groups or periods, not individuals. The literary historian will speak of the Age of Louis XIV, not the Age of Corneille or Racine or Molière or whomever, and in a sense the tradition as a whole is more significant than any individual in it (Curtius 1932, p. 189). The many competing literary groups encourage intellectual quality and simultaneously discourage lyric effusions, indeed extremes of any kind. Not surprisingly, many observers have contrasted the French talent for prose with the English gift for poetry—the English poet A. E. Housman once even asked André Gide why French literature had no poets at all!

American elites give no such dramatic coherence to American literary culture. There are, of course, important connections between literature and elites, but they are looser and less conspicuous than in France. Despite the undeniable importance of New York, elites, like writers, are dispersed. Moreover, no entrenched tradition of governmental patronage gave the American writer a sense of importance and of belonging to the center of society rather than being relegated to the periphery. To be sure, there are instances of governmental patronage of letters (Charvat 1968) and in 19th-century Boston the writer was "both the favourite son and the spokesman of the city of culture" (Green 1967, p. 79) where, in the words of one contemporary, "to be known as an able writer is to have the choicest society opened to you" (Green 1967, p. 72). But Yankee Boston is usually considered something of an exception to American mores, if not an aberration. More recognizable, more "American," are the patterns of affiliation characteristic of the Chicago literary world where newspapers exercised much more influence than genteel circles (Duncan 1961).

The many French literary groups which serve as buffers between the writer and mass culture are also fewer or more scattered in the United States. Consequently, the American writer is more at the mercy of popular culture (which may have something to do with the virulence of the so-called mass culture controversy of the 1950s for which there was no French equivalent). The gulf between "serious" (elite by one criterion or another) and popular literature also seems less broad in the United States than in France, perhaps simply because intellectual (and other) constraints are fewer and elites dispersed.

American support for literature has come almost entirely from the private sector, from the new commercial elites of self-made men instead of from an aristocracy which gravitated around the crown. Literature found support through foundations, prize associations, libraries, and eventually universities, which are the largest single source of American literary prizes (29.8%) (no French university sponsors a prize) and which have provided employment for writers ever since Longfellow was appointed to Bowdoin, then

Harvard. Because a French university chair requires the *doctorat d'état* (acquired with considerably more difficulty than a Ph.D.), even writers with advanced degrees like the *licence* or *agrégration,* when they taught, did so at the secondary level. Partly by virtue of this exclusion, partly because of a pervasive distrust of modern literature, the French university never welcomed even the brightest and best of contemporary writers or writing. Only within the past decade have *doctorat d'état* candidates been permitted to write theses on living writers. Creative writing courses, programs, even advanced degrees, have no French equivalent. It is the pluralism of American literature which often strikes French observers (among others), the multiple voices with which it has always spoken, the significance of regional and local differences, the openness, even in academic circles, to popular genres like the detective novel, the western, or science fiction. Because its elite connections are so conspicuous and so ingrained, and because the resulting literary consciousness is highly developed, French literary culture has seemed more of a piece and more elitist than American literary culture.

V

Clearly the centripetal tendencies of French society have reinforced those of French literary culture, while the centrifugal forces of American society have tended to deflect elite tendencies and to reinforce local over national standards. New York has never gone unchallenged, and that challenge has probably been amplified since the National Endowments for the Arts and the Humanities were established in 1965. Not only has there been a certain displacement of literary concerns to Washington, but more important still, the endowments have actively promoted state arts and humanities councils (Netzer 1978, pp. 89–93). The concentration of literary activities in Paris, which facilitates and almost requires personal contact among writers, has done much to sustain a sense of national literary community largely missing in the United States (see Bellow 1977).

Not even New York approaches the dominance Paris has for French letters.[15] Surely a good deal of the excitement writers have found there is generated by the sheer concentration of literary people and resources. Tele-

[15] Analysis of 1970 Census data for 25,376 employed authors showed that 17.3% resided in New York City, 9 1% in Los Angeles, 8.1% in Washington, D C , 2 9% in San Francisco, and 2.3% in Chicago. By state the range went from 19.97% of authors in New York and 17.55% in California to 5.13% in Maryland and 4.94% in Massachusetts (NEA 1977, pp. 47, 58). Since the "employed author" classification includes professional writers outside the traditional fine arts occupations (NEA 1977, p. 12) one should weight New York more heavily. Nevertheless, if New York dominates the literary scene, it does not stand alone or even in the majority. The selective *Directory of American Poets* of 1975 (Poets and Writers 1975)—a panel evaluates all applications for listings—indicates that 24.5% (330) of the 1,536 listed poets resided in New York City, 13% (201) in New York State outside New York City, 6% (92) in Massachusetts, and 2%–3% in a number of states (Ohio, North Carolina, Michigan, Illinois, Pennsylvania, Connecticut).

phone directories listed 455 bookstores in New York, 431 in London, 337 in Chicago, but 1,239 in Paris. There is a similar concentration of publishers. Telephone directories listed 584 book publishers in Manhattan, 432 in London, 154 in Chicago, but 1,124 in the Paris region, or 74% of the 1,514 publishers in France (289, or 19%, are in the fifth, sixth, and seventh *arrondissements* alone.) Of the 371 largest French publishers (those with sales over $20,000), 321 (87%) are in the Paris region, 302 (81%) in the city itself (Troubnikoff 1972, p. 123), whereas the whole of the Middle Atlantic region, according to the *Census of Manufactures* (U.S. Bureau of the Census 1971), has no more than 42% of all American publishers (which account for 55% of all sales).

Traditionally, a common social and educational background reinforced the connections of literary and other elites in France. Like French elites generally, literary elites came from existing elite milieus (Girard 1961, pp. 89–97; Escarpit 1964, pp. 44–47)[16] and received a common education preparing for the *baccalauréat* (Girard 1961, p. 104; Prost 1968, pp. 58, 245, 331–32). Unlike the American high school diploma received at about the same age, the *baccalauréat* was an elite degree: only 1.35% of boys between 11 and 17 attended a lycée or collège in 1850, 2.59% in 1900, 3.58% in 1930 (Isambert-Jamati 1970, p. 376).

The lycée program, which was determined in detail by the Ministry of Education, was particularly important in inculcating familiarity with and respect for literature. The philosophical and literary component was dominant; the plastic or musical arts received little attention (Ardagh 1968, p. 320; Isambert-Jamati 1970, pp. 380–81). French was taught through Latin models, and Latin, along with the general culture it signaled, both united elites and distinguished them from everyone else (Goblot [1925] 1967). The stress on development of a "correct" style, on effective oral and written expression through the imitation of literary models (Genette 1966), could not but convey the prestige of the acknowledged masters of such expression—writers. The diversity of the American educational system precluded any such common educational (and social) experience.

Political structures reinforced the elite orientation of French society and, by extension, French literary culture. Far more than their American counterparts and well beyond the ancien régime, French political institutions favored the dominance of national elites,[17] whereas the considerable auton-

[16] Of 623 French *gens de lettres* from 1300 to 1825, 26% came from the aristocracy, 30% from the magistrature (judiciary and public administration) Aristocrats declined from 50% of the total for the period 1330–1550 to 16% for 1776–1825 (Odin 1895). For government service in particular, two samples of 19th-century French and English writers by Escarpit (1964, p. 44) show that 16% of the French writers were diplomats or high administrative officials versus only 8% of the English.

[17] A brief trial of universal (manhood) suffage early in the First Republic (1792–99) was succeeded by the Napoleonic Consulate (1799–1804) and authoritarian (First) Em-

omy of American local and state governments supported strong regional identities and provoked suspicions of, rather than identification with, national elites. Especially where governmental interests were at stake, elite notions of any sort tended to be defused by a populism recurrent in American life. Jacksonian "democracy" effectively quashed John Quincy Adams's dream of a national academy of arts and sciences, and when Secretary of State Sumner made a similar proposal in the 1860s, it was at best an anomaly.

VI

The past quarter century has considerably modified these traditional patterns of support. Interest in and a certain sympathy for cultural activities have broadened and created pressures to which the changed and changing tenor of French as of American society has permitted and even forced response. The once very great distance between French and American literary cultures has diminished accordingly.

Modernization of French industry and the economic boom of the 1950s and 1960s could not but affect the production and marketing of books. The mergers of many small (and not so small) publishing houses effected a certain rationalization of commercial procedures, and the image of the writer as adventurer is gradually fading before that of the "professional" in the lay sense for whom specimen contracts, regular royalty statements and disbursements, and social security affiliation may well become the rule rather than the exception.

Democratization of French political institutions (the first direct popular election of the head of state occurred in 1962), often strident criticisms of the status quo coming from within the government as well as from without, the *événements* of 1968 with their dramatic contestation of cultural and other authority, and the coalitions on the Left joined to make the government more sensitive to such popular demands. In its cultural policy the Fifth Republic has sought actively to democratize cultural activities and to attenuate their ingrained elite associations André Malraux, France's first minister of culture, went so far as to speak of the cultural "rights" of all the French.

But it is the extension of education, of higher education in particular,

---

pire (1804–14). During the Bourbon Restoration (1814–30) suffrage was based on property taxes which made 1% of the adult male population eligible to vote. The July Monarchy (1840–48), while keeping the principle of tax-based eligibility, lowered the minimum to include 8% of adult males—and this after a revolution. In both regimes *eligibility for election required a still higher tax base. After the brief euphoria of the* Second Republic (1848–52), the Second Empire (1852–70) returned to autocracy, and even the Third Republic (1870–1940), which had universal manhood suffrage, remained fairly elitist in its early years. French women did not vote until 1945, a quarter of a century after American women.

that is the essential force in the rapprochement of the two literary cultures. In France the lycées and collèges, which in 1951 included only 8.28% of the 11–17 age cohort (male), expanded to 19.65% by 1964, with a mean annual growth rate of 8.6% in 1960–64 alone (Isambert-Jamati 1970, pp. 376–77). Higher education has followed suit with an average annual compound growth rate in enrollment between 1960 and 1970 of 11.2% in France and 8.3% in the United States (OECD 1974, p. 32). The estimated change in average per capita educational attainment of the population aged 25–64 in France was 3.2% from 1948–58, 3.1% in 1958–68, and an estimated 6.1% for 1968–78; in the United States 2.7% for 1950–60, 3.1% for 1960–70, and an estimated 5.6% for 1970–80 (OECD 1974, pp. 60–62).

In sheer numbers alone students constitute a formidable reading public, not only for textbooks but for works of all kinds. From 282,400 full-time students in higher education in 1960 the French student population almost tripled by 1970 to 846,200; the American student population in the same years more than doubled from 2,466,000 to 5,490,000 (OECD 1974, p. 20). And these figures do not count the part-time students, the habitués and hangers-on in university communities from Berkeley to the Left Bank.

The expanded reading publics have affected publishing in both countries. From 1880 to 1960 the number of titles published annually grew in the United States at an average rate of 2.5% and attained an annual growth rate of 7.8% from 1960 to 1969 (Hokkanen 1969, p. 66). French publishing also increased its production at an annual rate of 6.6% between 1960 and 1970 (Troubnikoff 1972, p. 107). Given the consistent correlation of reading with education, this growth in the educated population bodes well for the recruitment of readers in years to come. However, these "new" readers, increasingly from lower socioeconomic groups, will not resemble the traditional lettered minority. The elite educational patterns once so salient in France will be greatly altered as will reading preferences. Even now popular literature seems in a stage of redefinition as semirespectable. Literary prizes already exist for detective novels and science fiction, and comics cannot be far behind.

In the United States an extension of sympathy, even a demand, for cultural activities coupled with strong support within the federal government culminated in the establishment of the National Endowments for the Arts and the Humanities (20% of NEH funds can be considered support of the arts [Netzer 1978, p. 76]). From an initial budget of $2.5 million, in just a decade each endowment increased its appropriation to $74.7 million. (Appropriations for the 1979 fiscal year are $139.6 million.) During the same period, and partly under NEA and NEH sponsorship, state councils for the arts and the humanities also increased in numbers and in funding (Netzer 1978, pp. 89–93). Even the U.S. Conference of Mayors has an Arts Task Force. However, because American support is recent and also because Ameri-

can legislators are highly sensitive to charges of elitism, more care is taken (and is needed) to justify appropriations and even grants than is the case in France where centuries of funding legitimate government subsidies of the arts.[18]

This notable rapprochement notwithstanding, French and American literary cultures are distinct. Many basic social structures of French and American society remain familiar if not intact. Paris is still the center of French society and of French literary culture, still the locus of French elites. Despite efforts to decentralize funding, the French government continues to dominate patronage of literature as of the arts generally, while the American government walks gingerly on the same path. Education in France remains markedly more elitist than in the United States, primarily because recruitment is more selective: 19.5% of 19-year-olds, 16.2% of 20-year-olds in France were enrolled in institutions of higher education in 1970; the analogous figures for American students are 40.9% and 35.4% (OECD 1974, p. 30).

VII

Not all of society or all of literature equally influences literary culture. Originally the province of an elite, literature, like the arts more generally, has remained dependent on elite support even, or especially, in the era of the mass market. And it is the diverse ties between literary activities and elites which inform, circumscribe, and define literary culture. If literary matters in France often turn into public affairs, this extraliterary resonance is largely a function of the multiple literary associations of national and particularly governmental elites. These ties and associations sustain a generalized elite concern for literary activities, the marked intensity of which contrasts sharply with the more diffuse and, as it were, more private literary ties of American elites.

It is undoubtedly true that the elite ties of literature in France are not unique. Some of the features ascribed to French literary culture are perhaps as European as they are French. Such similarity is scarcely surprising. After all, France shares much with England, Germany, perhaps Western European countries generally: a common legacy of aristocratic patronage of literature and the arts, the congruence of literary and political elites characteristic of

[18] Netzer (1978, pp. 73, 75) estimates that 43% of all NEA grants from 1971–74 were designed to increase the availability of the arts. Without the matching grants from private contributors, the proportion for 1975 and 1976 would be 50%. Similarly, after 1974 NEH shifted its emphasis to its Division of Public Programs, which stressed diffusion to a broad audience. As their budgets increase the endowments become increasingly embroiled in congressional politics (NEA has been honored by one of Senator Proxmire's Golden Fleece awards), hence increasingly subject to local and regional requirements as well as to demands for the "popularization" of the arts.

aristocratic societies, elite educational patterns that stressed literature and the classics (Armstrong 1974, chap. 9), important involvement of current governments in subsidizing literature and the arts. It is only to be expected that literary cultures in these countries should resemble the French pattern more than the American.

But each of these countries is also heir to a very specific set of social, political, economic, literary, and linguistic traditions which necessarily create distinct, and to some extent unique, relations between literary activities and elites. The decentralization of German society and culture, for example, recalls the United States more than France. Likewise, the important hiatus in England between the end of aristocratic patronage of letters in the 18th century and the commencement of government support in the 20th century created circumstances closer to the American than to the French context. The literary cultures sustained by these and other indigenous traditions cannot but be distinct, nearer the French than the American model no doubt, but not identical with it.

REFERENCES

Ardagh, John. 1968 *The New French Revolution*. New York. Harper & Row

Armstrong, John A. 1974 *The European Administrative Elite* Princeton, N.J : Princeton University Press.

Arnold, Matthew (1864) 1962 "The Literary Influence of Academies," *Complete Prose Works*. Vol. 3. Ann Arbor: University of Michigan Press

Aron, Jean-Paul, ed. 1975. *Qu'est-ce que la culture française* Paris· Denoël-Gonthier.

Bellow, Saul. 1977. "Writers and Literature in American Society." Pp 172–96 in *Culture and Its Creators*, edited by J. Ben-David and T N. Clark. Chicago: University of Chicago Press.

Bengtsson, Sverker. 1968. *La Défense organisée de la langue française* Uppsala· Almquist & Wiksells.

Bourdieu, Pierre, with Luc Boltanski 1975 "Le Fétichisme de la langue," *Actes de la Recherche en sciences sociales*, no. 4 (July), pp. 2–32.

Charvat, William. 1968. *The Profession of Authorship in America*. Columbus: Ohio State University Press.

Clark, Priscilla P. 1976 "Deux Types de subventions: L'Assistance aux écrivains en France et aux Etats-Unis." *Bibliographie de la France* 24 (June)· 1232–42.

———. 1977. "Styles of Subsidy· Support for Writers in France and the United States." *French Review* 50 (March): 543–49.

———. 1979. "Indicateurs socio-culturels et culture littéraire." *Revue de la Société Tocqueville*, in press.

Clark, Priscilla P., and T. N. Clark. 1974. "Stratification and Culture: The Position and Roles of Intellectual and Other Elites in France and the United States." Mimeographed. Available from author.

———. 1977. "Patrons, Publishers, and Prizes· The Writer's Estate in France " Pp. 197–225 in *Culture and Its Creators*, edited by J. Ben-David and T. N. Clark. Chicago: University of Chicago Press.

Curtius, E. R. 1932 *Essai sur la France*. Paris: Grasset.

Dumazedier, Joffre, and Jean Hassenforder. 1962 "Eléments pour une sociologie comparée de la production, de la diffusion et de l'utilisation du livre," *Bibliographie de la France*, no. 24 (June 15), pp. 1–100.

Duncan, H. D. 1961. *The Rise of Chicago as a Literary Center 1885–1920*. Totowa, N.J.: Bedminster Press.
Ennis, Philip H. 1972. "Who Reads?" Pp. 37–56 in *The Metropolitan Library*, edited by Ralph Conant and Kathleen Molz. Cambridge, Mass.: M.I.T. Press.
Escarpit, Robert. 1964. *La Sociologie de la littérature*. Paris: Presses Universitaires de France.
Le Figaro (April 12–18, 1976).
L'Express (September 16–22, 1974).
Frémy, Dominique, and Michèle Frémy, eds. 1976. *Quid*. Paris: Laffont.
Genette, Gérard. 1966 "Enseignement et rhétorique du XXe siècle." *Annales—Economies, Sociétés, Civilisations* 21, no. 2, pp. 293–305.
Gerbner, George. 1973. "Cultural Indicators: The Third Voice." Pp. 555–73 in *Communications Technology and Social Policy*, edited by George Gerbner, L. Gross, and W. Melody. New York: Wiley.
Gerbner, George, L Gross, M. Jackson-Beeck, S. Jeffries-Fox, and N Signorelli. 1978. "Cultural Indicators: Violence Profile No. 9." *Journal of Communication* 28 (Summer): 176–206
Girard, Alain. 1961. *La Réussite sociale en France*. Paris: Presses Universitaires de France.
Goblot, Edmond. (1925) 1967. *La Barrière et le niveau: Essai sur la bourgeoisie française moderne*. Paris: Presses Universitaires de France.
Gouillou, André. 1975 *Le Book-Business ou l'édition française contre la lecture populaire*. Paris: Tema-éditions.
Green, Martin. 1967. *The Problem of Boston: Some Readings in Cultural History*. New York: Norton.
Hillairet, Jacques. 1963. *Dictionnaire historique des rues de Paris*. Vol. 2, 4th ed. Paris: Editions de Minuit.
Hokkanen, Dorothy. 1969. "American Book Title Output, 1880–1969." Pp. 65–70 in *Bowker Annual*. New York: Bowker.
Isambert-Jamati, Viviane. 1970. *Crises de la société crises de l'enseignement*. Paris: Presses Universitaires de France.
Kroeber, A. L. (1944) 1963. *Configurations of Culture Growth*. Berkeley: University of California Press.
Lottman, Herbert R. 1975. "The French: They are Hooked on La Télé." *New York Times* (April 20).
Lough, John. 1978. *Writer and Public in France from the Middle Ages to the Present Day*. Oxford: Clarendon.
National Endowment for the Arts (NEA). 1977. "Where Artists Live: 1970." Research Division Report no. 5 (October).
Netzer, Dick. 1978. *The Subsidized Muse: Public Support for the Arts in the United States*. Cambridge: Cambridge University Press.
New York City TV Guide (June 12–18, 1976).
Odin, Alfred. 1895. *Genèse des grands hommes: Gens de lettres français modernes*. Pp. 281–92 in Girard 1961.
Organization for Economic Cooperation and Development (OECD). 1974. *Educational Statistics Yearbook*. Paris: OECD.
Poets and Writers, Inc., ed. 1975. *Directory of American Poets*. New York: Poets & Writers.
Prost, Antoine. 1968. *L'Enseignement en France 1800–1967*. Paris: Armand Colin.
Renan, Ernest. 1860. "L'Académie Française." Pp. 335–51 in *Essais de morale et de critique*. Paris: Lévy.
Stein, Gertrude. (1940) 1970. *Paris France*. New York: Liveright.
Szalai, Alexander, ed. 1972. *The Use of Time*. The Hague: Mouton.
Thibaudet, Albert. 1940. "Pour la géographie littéraire." Pp. 136–43 in *Réflexions sur la littérature*. Vol. 2. Paris: Gallimard.
"Time" *National Subscribers Study*. 1976. New York: Time-Life.

de Tocqueville, Alexis. (1840) 1958. *Democracy in America*. Vol. 2. New York: Vintage Books.

Treiman, Donald J. 1977. *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

Troubnikoff, Marie. 1972 "Les Données numériques " Pp. 95–129 in *Le Livre français*, edited by J. Cain, R. Escarpit, H.-J. Martin. Paris: Imprimerie Nationale.

UNESCO. 1973, 1975. *Statistical Yearbook*. Paris· UNESCO.

United Nations. *Statistical Yearbook*. Various issues. New York: United Nations

U.S. Bureau of the Census. 1971. *Census of Manufactures 1967*. Vol. 1. Washington, D.C.: Government Printing Office.

Wilson, Edmund. 1971 *Upstate*. New York: Farrar, Straus & Giroux.

A Brief Guide to the Causal Analysis of Data from Surveys[1]

Leo A. Goodman
University of Chicago

Appropriate methods for the causal analysis of certain kinds of systems of quantitative variables have been available for some time now, and they provide useful tools for the research worker in sociology, and in other fields as well, who may be concerned with the study of quantitative variables. Appropriate methods for the causal analysis of corresponding kinds of systems of qualitative variables have been introduced more recently, and they provide useful tools for the research worker who may be concerned with the study of qualitative variables. In order to facilitate the application of these more recently developed methods to the causal analysis of qualitative variables, further exposition of these methods is presented in this note. One of the contributions of this additional exposition is the correction of incorrect statements on this topic by H. T. Reynolds.

The utility of the methods discussed herein will be illustrated by applying them to data obtained in two well-known surveys. We shall reexamine some data on student attitudes analyzed earlier by Coleman (1964) and Goodman (1973*a*) and also some data on voting intention analyzed earlier by Lazarsfeld (1948), Lipset et al. (1954), Campbell (1963), Lazarsfeld, Berelson, and Gaudet (1968), Boudon (1968), Lazarsfeld (1970), and Goodman (1973*a*). These data are reproduced here as tables 1 and 2, respectively. Table 1 cross-classifies the responses of 3,398 schoolboys, each interviewed at two successive points in time, with respect to questions about their self-perceived membership in the "leading crowd" (in it or out of it) and their attitude concerning it (whether they think that membership in it does not require going against one's principles sometimes or whether they think that it does); and table 2 cross-classifies the responses of 266 people, each interviewed at two successive points in time, with respect to their vote intention (Republican vs. Democrat) and their opinion of a particular candidate (for or against the Republican candidate). In the present note, special attention will be given to the causal analysis of these data, using rather simple systems of equations which can be represented, in part, by

[1] The research reported in this note was supported in part by research contract SOC 76-80389 from the Division of the Social Sciences of the National Science Foundation. For helpful comments, the author is indebted to C. C. Clogg and O. D. Duncan.

TABLE 1

OBSERVED CROSS-CLASSIFICATION OF 3,398 SCHOOLBOYS, IN INTERVIEWS AT TWO SUCCESSIVE POINTS IN TIME, WITH RESPECT TO TWO DICHOTOMOUS VARIABLES*

		Second Interview			
Membership		+	+	−	−
Attitude		+	−	+	−
First Interview					
Membership	Attitude				
+	+	458	140	110	49
+	−	171	182	56	87
−	+	184	75	531	281
−	−	85	97	338	554

SOURCE.—Coleman 1964.

NOTE.—With respect to (self-perceived) membership in the "leading crowd," being in it is denoted + and being out of it is denoted −; with respect to attitude concerning the "leading crowd," a favorable attitude is denoted + and an unfavorable attitude is denoted −. Membership and attitude will be denoted by the letters *A* and *B*, respectively, for the first interview, and by the letters *C* and *D*, respectively, for the second interview

* Variables = (1) self-perceived membership in the "leading crowd", (2) favorableness of attitude concerning "leading crowd."

TABLE 2

OBSERVED CROSS-CLASSIFICATION OF 266 PEOPLE, IN INTERVIEWS AT TWO SUCCESSIVE POINTS IN TIME, WITH RESPECT TO TWO DICHOTOMOUS VARIABLES*

		Second Interview			
Vote Intention		+	+	−	−
Candidate Opinion		+	−	+	−
First Interview					
Vote Intention	Candidate Opinion				
+	+	129	3	1	2
+	−	11	23	0	1
−	+	1	0	12	11
−	−	1	1	2	68

SOURCES.—Lazarsfeld 1948; Lipset et al 1954, and Lazarsfeld, Berelson, and Gaudet 1968, among others.

NOTE.—With respect to vote intention, + and − denote Republican and Democrat, respectively, with respect to opinion of a candidate, a favorable opinion is denoted + and an unfavorable opinion is denoted −. Vote intention and candidate opinion will be denoted by the letters *A* and *B*, respectively, for the first interview; and by the letters *C* and *D*, respectively, for the second interview.

* Variables = (1) vote intention, (2) candidate opinion.

path diagrams of the kind displayed in figures 1 and 2, pertaining to tables 1 and 2, respectively.[2]

In figure 1, variables *A* and *B* pertain to self-perceived membership in the "leading crowd" and attitude concerning the "leading crowd," respectively, in the first interview; *C* and *D* pertain to the corresponding variables in the second interview. In figure 2, variables *A* and *B* pertain to vote intention and candidate opinion, respectively, in the first interview; *C* and *D* denote the corresponding variables in the second interview. Figures 1 and 2 pertain to the situation in which variables *A* and *B* are viewed as being prior to *C* (in the sense that the observed level of *A* and *B* might affect the odds pertaining to *C*); and variables *A*, *B*, and *C* are viewed as being prior to *D* (in the sense that the observed level of *A*, *B*, and *C* might affect the odds pertaining to *D*). In these figures, the prior variables *A* and *B* can be viewed as being mutually related to each other, or as being variables that affect each other.[3] In figure 1, there is a main effect of variable *A* on *C*

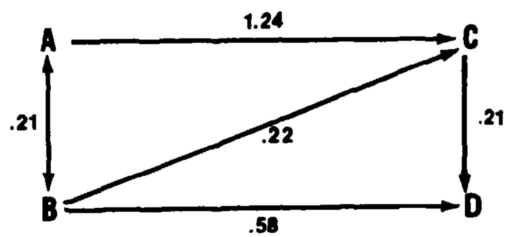


FIG. 1.—Path diagram for the system of eqq. (1a), (1b), (3), and (6), applied to table 1

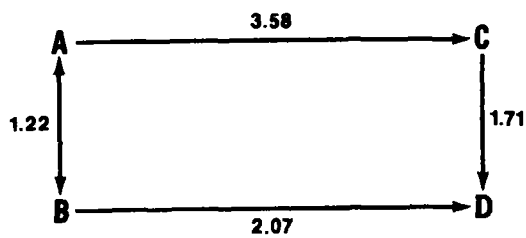


FIG. 2.—Path diagram for the system of eqq. (1a), (1b), (4), and (6), applied to table 2

[2] Figs. 1 and 2 display path diagrams corresponding to certain recursive systems for the qualitative variables *A*, *B*, *C*, and *D*. The remarks on causal analysis in Reynolds (1977) were limited to the analysis of recursive systems, and one of the contributions of the present note will be the correction of incorrect statements on this topic in that article. Special attention will be devoted here to the analysis of recursive systems, but the analysis of nonrecursive systems will also be commented on later herein.

[3] The double-headed arrow between variables *A* and *B* in figs. 1 and 2 can represent either (*a*) the mutual relationship between these two prior variables or (*b*) the effects both of variable *A* on *B* and of variable *B* on *A*. (When the prior variables *A* and *B*

and a main effect of B on C, and there is no interaction between these effects on C; in addition, there is no main effect of A on D, there is a main effect of B on D and a main effect of C on D, and there is no interaction between these effects on D. Figure 2 includes the same kinds of effects as in figure 1 except that there is no effect of variable B on C in figure 2.

To determine whether a specific causal model is congruent with the observed data, the observed frequencies in the table are compared with the expected frequencies estimated under the model, and an assessment is made using the usual goodness-of-fit χ^2 statistic or the corresponding χ^2 statistic based upon the likelihood ratio.[4] When the causal model represented by figure 1 is applied to table 1, the likelihood-ratio χ^2 turns out to be 4.06 with 6 degrees of freedom; and when the causal model represented by figure 2 is applied to table 2, the corresponding χ^2 turns out to be 1.61 with 7 degrees of freedom. Thus, these causal models fit the observed data very well indeed. Since each of these causal models consists of three component models (namely, model H_1 pertaining to the relationship between variables A and B; model H_2 pertaining to variable C with A and B viewed as being prior to it; and model H_3 pertaining to variable D with variables A, B, and C viewed as being prior to it), each of the χ^2 values reported above can be partitioned into its three component χ^2s. This partitioning is summarized in table 3.

Model H_1 in table 3 pertains to the mutual relationship between variables A and B in the two-way table $\{AB\}$. Here the saturated model for table $\{AB\}$ could be considered, or we could consider instead various unsaturated models for table $\{AB\}$. The saturated model can be described by the following formulae for the odds $\Omega_i^{A\bar{B}}$ pertaining to variable B and the odds $\Omega_j^{B\bar{A}}$

---

are viewed as being mutually related, the association between these variables is considered to be due to unobserved sources, which are not included explicitly in the model, in addition to the possible influence of variable A on B and/or variable B on A; whereas when the prior variables are viewed as variables that affect each other, the relationship between the variables is considered to be due to the effects of A on B and of B on A) If variable A had been viewed as being prior to variable B (in the sense that the observed level of A might affect the odds pertaining to B), then a single-headed arrow would have been pointed from A to B; and if variable B had been viewed as being prior to variable A (in the sense that the observed level of B might affect the odds pertaining to A), then a single-headed arrow would have been pointed from B to A. The results presented herein can be directly applied both to the case in which a single-headed arrow is pointed from A to B and to the case in which a single-headed arrow is pointed from B to A, as well as to the case in which there is a double-headed arrow between A and B.

[4] For the sake of brevity and simplicity, this note will not include numerical results for both the goodness-of-fit χ^2 and the likelihood-ratio χ^2. Numerical results for only the latter χ^2 will be presented here. The interested reader can find numerical results for both χ^2 statistics in Goodman (1973*a*, 1973*b*), and he can also find there the appropriate formulae for calculating these quantities as well as the appropriate formulae for calculating the expected frequencies estimated under the causal model.

pertaining to variable A:

$$\Omega_i^{A\bar{B}} = \gamma^{\bar{B}}\gamma_i^{A\bar{B}}\,, \tag{1a}$$

$$\Omega_j^{B\bar{A}} = \gamma^{\bar{A}}\gamma_j^{B\bar{A}}\,; \tag{1b}$$

and the unsaturated models for these odds could be described in similar terms (see, e.g., Goodman 1973*a*, 1973*b*). In both figures 1 and 2, model H_1 for table $\{AB\}$ was taken to be the saturated model.

Model H_2 in table 3 pertains to variable C with A and B viewed as being prior to it. Here the saturated model for the three-way table $\{ABC\}$ could be considered, or we could consider instead certain kinds of unsaturated models for table $\{ABC\}$. The saturated model can be described by the following formula for the odds $\Omega_{ij}^{AB\bar{C}}$ pertaining to variable C:

$$\Omega_{ij}^{AB\bar{C}} = \gamma^{\bar{C}}\gamma_i^{A\bar{C}}\gamma_j^{B\bar{C}}\gamma_{ij}^{AB\bar{C}}\,; \tag{2}$$

and the unsaturated models that we might consider for the odds $\Omega_{ij}^{AB\bar{C}}$ could be described in similar terms. In figure 1, model H_2 for table $\{ABC\}$ was taken to be

$$\Omega_{ij}^{AB\bar{C}} = \gamma^{\bar{C}}\gamma_i^{A\bar{C}}\gamma_j^{B\bar{C}}\,, \tag{3}$$

and in figure 2, model H_2 for table $\{ABC\}$ was taken to be

$$\Omega_{ij}^{AB\bar{C}} = \gamma^{\bar{C}}\gamma_i^{A\bar{C}}\,. \tag{4}$$

Model H_3 in table 3 pertains to variable D with variables A, B, and C viewed as being prior to it. Here the saturated model for the four-way

TABLE 3

PARTITIONING OF TOTAL VARIATION UNEXPLAINED BY PATH DIAGRAM IN FIGURES 1 AND 2 APPLIED TO TABLES 1 AND 2, RESPECTIVELY

Source of Variation	df	Likelihood Ratio χ^2
	Figure 1 Applied to Table 1	
1. Variation unexplained by H_1 in $\{AB\}$ . . .	0	0 00
2. Variation unexplained by H_2 in $\{ABC\}$ . .	1	0 02
3. Variation unexplained by H_3 in $\{ABCD\}$.	5	4 04
Total variation unexplained by fig. 1 . .	6	4 06
	Figure 2 Applied to Table 2	
1. Variation unexplained by H_1 in $\{AB\}$	0	0.00
2. Variation unexplained by H_2 in $\{ABC\}$.	2	0 15
3. Variation unexplained by H_3 in $\{ABCD\}$ . .	5	1 45
Total variation unexplained by fig. 2 .	7	1.61*

* The sum is 1.61, rather than 1.60, when more significant digits are used than are reported here.

table $\{ABCD\}$ could be considered, or we could consider instead certain kinds of unsaturated models for table $\{ABCD\}$. The saturated model can be described by the following formula for the odds $\Omega_{ijk}^{ABC\bar{D}}$ pertaining to variable D:

$$\Omega_{ijk}^{ABC\bar{D}} = \gamma^{\bar{D}}\gamma_i^{A\bar{D}}\gamma_j^{B\bar{D}}\gamma_k^{C\bar{D}}\gamma_{ij}^{AB\bar{D}}\gamma_{ik}^{AC\bar{D}}\gamma_{jk}^{BC\bar{D}}\gamma_{ijk}^{ABC\bar{D}} ; \qquad (5)$$

and the unsaturated models that we might consider for the odds $\Omega_{ijk}^{ABC\bar{D}}$ could be described in similar terms. In both figures 1 and 2, model H_3 for table $\{ABCD\}$ was taken to be

$$\Omega_{ijk}^{ABC\bar{D}} = \gamma^{\bar{D}}\gamma_j^{B\bar{D}}\gamma_k^{C\bar{D}} . \qquad (6)$$

The overall causal models presented above, which are described, in part, by the corresponding path diagrams, make explicit what is the assumed order of priority of the variables in the model; and this assumed order of priority will determine which are the relevant multiplicative models to be considered. Thus, for the order of priority assumed above, the relevant multiplicative models to be considered are the saturated models (1a), (1b), (2), (5), and corresponding unsaturated models. These models pertain to the odds $\Omega_i^{A\bar{B}}$, $\Omega_j^{B\bar{A}}$, $\Omega_{ij}^{AB\bar{C}}$, and $\Omega_{ijk}^{ABC\bar{D}}$, respectively; and each of the corresponding multiplicative models for the odds can be replaced by the equivalent additive (logit) model for the log-odds (see, e.g., Goodman 1972*a*, 1973*a*).

The numerical magnitude associated with each of the arrows in figures 1 and 2 is the magnitude of the corresponding effect as estimated with the appropriate logit model obtained when the multiplicative models presented above for the odds are replaced by the equivalent additive models for the log-odds. These numerical magnitudes are of substantive interest in understanding the phenomena analyzed with the data in tables 1 and 2. These quantities can be interpreted in a manner similar to the interpretation of the effects in the usual additive model used in the analysis of variance and to the interpretation of the slopes in the usual multiple-regression model (see, e.g., Goodman 1975). Some further comments on these magnitudes will be made later herein.

CAUSAL ANALYSIS AND THE ANALYSIS OF MARGINAL TABLES

In the preceding comments on the causal analysis pertaining to the path diagrams in figures 1 and 2 applied to tables 1 and 2, respectively, we took note of the fact that the assumed order of priority of the variables, which is made explicit in the overall causal model, will determine which multiplicative models (or equivalent additive logit models) are relevant. After this determination of the relevant models has been made, these models would then be applied to the appropriate sets of data. Thus, for example, the

saturated models (1a) and (1b), and corresponding unsaturated models, would be applied to the two-way table $\{AB\}$; the saturated model (2), and corresponding unsaturated models, would be applied to the three-way table $\{ABC\}$; the saturated model (5), and corresponding unsaturated models, would be applied to table $\{ABCD\}$.

The assumed order of priority of the variables determined which models were relevant, and the models in turn determined which tables were relevant.[5] Of course, if the overall causal model under consideration is changed in ways that also change the assumed order of priority of the variables, then there will also be a change in the models that are relevant and in the tables that are relevant (i.e., in the tables to which these models are to be applied).

In the preceding discussion, with the particular order of priority assumed there, we were led to an analysis in turn of the tables $\{AB\}$, $\{ABC\}$, and $\{ABCD\}$, using the relevant multiplicative models for the corresponding odds or the equivalent additive logit models for the corresponding log-odds. There was no need in that discussion to consider other more general log-linear models applied to these tables.[6] Indeed, the more general log-linear models (or the equivalent more general multiplicative models), applied to these tables, would not be particularly relevant in the analysis of recursive systems of the kind we have been considering here.[7] However, the overall causal model pertaining to the analysis of a recursive system, which can be described either in terms of a set of additive logit models for the log-odds or in terms of a corresponding set of equivalent multiplicative models for the odds (e.g., the set of models [1a], [1b], [3], and [6] used to describe the causal model in fig. 1), sometimes turns out to be equivalent to a single log-linear model (or a corresponding single [more general] multiplicative model) applied to the full table $\{ABCD\}$. This was the case for the causal model in figure 1, but it was not the case for the causal model in figure 2. (For further details on these matters, see, e.g., Goodman [1973*a*, 1973*b*].)

Note that the appropriate methods of causal analysis presented above led to an analysis in turn of tables in which certain specified variables were collapsed. In the analysis of the saturated models (1a) and (1b), and corresponding unsaturated models, since table $\{AB\}$ was used, variables C and D were collapsed; and in the analysis of the saturated model (2), and cor-

[5] We can view the assumed order of priority of the variables as determining which models are relevant and also which tables are relevant.

[6] The additive logit model for the log-odds can be viewed as a special case of more general log-linear models (see, e.g., Bishop 1969; Goodman 1970, 1971). Similarly, the multiplicative model for the odds can be viewed as a special case of a more general kind of multiplicative model (see, e.g , Goodman 1972*a*, 1972*b*, 1973*a*).

[7] The more general log-linear models (and the equivalent more general multiplicative models) are relevant in the analysis of nonrecursive systems. Some further comments on these more general models and on the analysis of nonrecursive systems appear later herein.

responding unsaturated models, since table {*ABC*} was used, variable *D* was collapsed. In the analysis of the saturated models (1a) and (1b), and corresponding unsaturated models, the existence or nonexistence of the possible two-factor effect (*AB*) between variables *A* and *B* in the two-way table {*AB*} is relevant; but the existence or nonexistence of a corresponding possible two-factor effect in the four-way table {*ABCD*} is not relevant when variables *A* and *B* are viewed as being prior to variables *C* and *D*. Similarly, in the analysis of the saturated model (2), and corresponding unsaturated models, the existence or nonexistence of the possible three-factor effect (*ABC*) between variables *A*, *B*, and *C* in the three-way table {*ABC*} is relevant; but the existence or nonexistence of a corresponding possible three-factor effect in the four-way table {*ABCD*} is not relevant when variables *A*, *B*, and *C* are viewed as being prior to variable *D*.

The remarks on this general topic in Reynolds (1977) are incorrect in various respects. Let us now consider some of the errors.[8]

SOME ERRORS

1. Irrelevant Criticism

Reynolds (1977) seems to be criticizing the methods of analysis presented in Goodman (1973*a*) on the grounds that, when these methods are used in the analysis of a specified overall causal model (with a specified order of priority for the variables in the causal model), and the relevant component models for the odds (or the log-odds) are applied to the relevant tables (e.g., tables {*AB*}, {*ABC*}, and {*ABCD*}, with the particular order of priority considered herein), then some of the effects that are included in the relevant component models may correspond to effects that could be omitted from a log-linear model applied to the full table. In particular, Reynolds seems to find fault with these methods when they are used in the analysis of the recursive system represented by figure 1 above (fig. 10 in Goodman 1973*a*) applied to the data in table 1, because these methods lead to the inclusion of the two-factor effect (*AB*) between variables *A* and *B* in the relevant component model pertaining to the two-way table {*AB*}, while a corresponding two-factor effect in a log-linear model applied to the four-way table may be less pronounced (or absent). From the results presented near the end of the preceding section, we see that Reynolds's supposed criticism is incorrect, since the existence or nonexistence of the two-factor effect (*AB*) in the two-way table {*AB*} is relevant in the causal analysis of the recursive system represented by figure 1, whereas the existence or nonexistence of a corresponding effect in the four-way table would not be relevant.

For some other causal models which are very different from the ones

[8] For additional comments on some of the errors, see Gillespie (1978).

considered here,[9] the existence or nonexistence of a two-factor effect (AB) in the four-way table would have some relevance; and these particular models have been discussed in Goodman (1973*a*, 1974).[10]

2. Incorrect and/or Irrelevant Simulation

Reynolds also seems to be criticizing the methods of analysis presented in Goodman (1973*a*) on the grounds that these methods would fail to detect the structure in a recursive system of *quantitative* variables when these variables have been dichotomized. But the structure of the system of dichotomized variables will usually be very different from the structure of the corresponding quantitative variables.[11] The methods of analysis presented in Goodman (1973*a*) and in this article were developed for the analysis of the structure of a system of dichotomous or polytomous variables; they were *not* developed for the analysis of the structure of a system of quantitative variables. (This is true also of the methods of analysis presented in Goodman's articles on related topics; see, e.g., his papers reprinted in Goodman [1978].) In Reynolds's attempt to see whether the methods described in Goodman (1973*a*) could detect the known underlying structure in a simulated recursive system, he incorrectly simulated a system of *quantitative* variables (which he then dichotomized) when he should instead have simulated a system of *qualitative* variables. Reynolds's simulation, presented in his section discussing Goodman's methods for the analysis of recursive systems of qualitative variables, is thus incorrect and/or irrelevant.

In addition to the simulation mentioned above, Reynolds also presented a simulation pertaining to the analysis of conditional independence. Here too he simulated a system of *quantitative* variables (which he then dichotomized and polytomized) instead of a system of *qualitative* variables. But none of Goodman's articles have suggested that the appropriate methods developed for the analysis of *qualitative* variables (e.g., methods developed for the analysis of conditional independence models, logit models, log-linear models, etc.) should be applied to arbitrarily dichotomized (or polytomized) *quantitative* variables. Indeed, the appropriate methods for

[9] In the causal models considered here variables A and B were assumed to be prior to variables C and D, whereas in the other causal models variables C and D would have to be assumed to be prior to variables A and B. Since variables A and B pertain to responses in a first interview, and variables C and D pertain to responses in a second interview, the latter causal models would have to involve apparent time reversal. For a discussion of related matters, see Goodman (1973*a*, 1974).

[10] See, e.g., figs. 3, 4, 8, and 9 in Goodman (1973*a*); fig 6 in Goodman (1974); and the discussion of models H''_3 and H_{13} in the latter article.

[11] For further comments on the difference between the structure of the system of dichotomized variables and the structure of the corresponding quantitative variables, see the Appendix below.

the analysis of *qualitative* variables should not be applied in this inappropriate way.[12]

3. Incorrect Invocation of Theorem on Collapsing Variables

In Reynolds's section on the analysis of recursive systems, he attempts to explain why, in the analysis of the three-way marginal table $\{ABC\}$ obtained with his simulated data, a two-factor effect (BC) was needed, and he incorrectly invokes Bishop, Fienberg, and Holland's (1975) theorem here. Reynolds states that "certainly we know from Bishop et al.'s theorem that such a term might be expected to appear. Analysis of the $\{ABC\}$ marginal table . . . almost guarantees that we will find such a link" (p. 139). This statement is incorrect, since (*a*) the term appeared in the analysis of the simulated data because the simulation was done incorrectly (see comment on this at the beginning of the preceding section), (*b*) if the simulation had been done correctly the term would not have appeared, (*c*) Bishop et al.'s theorem does *not* imply that such a term might be expected to appear, and (*d*) the analysis of the $\{ABC\}$ marginal table would *not* even "almost guarantee" that there would be such a link.[13]

Similar incorrect invocations also appear in other sections in Reynolds's article. For example, in his section on collapsing variables for the sake of convenience, for the particular data considered there Reynolds states that "when the variables have not been collapsed much simpler models become acceptable. . . . One can see how Bishop et al.'s theorem might account for these results. . . . By simply combining categories one may create artificial interactions which will give a misleading picture of the underlying relationships" (pp. 136–37). In a similar vein, Reynolds also states a little later that "collapsing a variable may require one to fit parameters that would otherwise be unnecessary" (pp. 139–40). In contrast to such statements, it should be noted that collapsing a variable, or combining particular categories of a variable, can lead either to the deletion of parameters that would otherwise be necessary,[14] or to the insertion of parameters that would otherwise be unnecessary, or to no change in the parameters.[15]

[12] For appropriate methods for determining whether particular categories of a polytomous variable can be combined, see, e.g., Duncan (1975).

[13] These brief remarks will be understood more fully by the reader after he has read the comments on the analysis of marginal tables in the final section herein.

[14] With respect to this effect of collapsing a variable, the analysis of table 1 in Goodman (1973*a*) (i.e., table 2 above) illustrated this in a dramatic way, and it will be discussed further in the final section below. (The reader is also referred to Goodman [1972*b*] for results that show in which direction the magnitude of a parameter will change when a dichotomous variable is collapsed [i e., deleted] or inserted) With respect to this effect of combining particular categories of a variable, the analysis in Duncan (1975) illustrated this well. The analysis of the effects of collapsing a variable (i.e., of combining *all* the categories of the variable into a single combined category) is quite different from the analysis of

4. Misleading Statements

In several places in his article, Reynolds makes statements that may be misleading and that are incorrect in a different way from the statements noted above. Here is one example.

In the section in which he criticizes collapsing the categories of a variable for the sake of convenience (e.g., dichotomizing a variable for convenience), he states that "in some cases [Goodman] has collapsed variables that originally had more than two categories. . . . The unwary reader may be left with the impression that the number of categories does not matter. He may feel that Goodman's . . . procedures permit him to manipulate data in any way he pleases . . ." (p. 135). This statement may be misleading, since (*a*) in none of my articles have I collapsed a variable that had more than two categories in the main source (or sources) that I used for comparative purposes,[16] and (*b*) these articles do *not* leave the impression that the number of categories does not matter and that the researcher is permitted to manipulate data in any way he pleases.[17]

the effects of combining particular categories of a variable (as long as all the categories are not combined into a single category). This distinction between the analysis of the effects of collapsing a variable and the analysis of the effects of combining particular categories of a variable is an important one. (Reynolds failed to take this distinction into account in his article.)

[15] This comment on the possible effects of collapsing (i.e., deleting) a variable, or combining particular categories of a variable, can also be stated in terms of the possible effects of inserting a variable, or decomposing a particular category of a variable into two or more subcategories. Thus, when a variable is inserted, or when a particular category of a variable is decomposed into two or more subcategories, this can lead either to the insertion of parameters that would otherwise be unnecessary, or to the deletion of parameters that would otherwise be necessary, or to no change in the parameters. The three possible effects listed above will be discussed more fully in the final section herein when the sign-rule in table 4 is considered.

[16] To illustrate the utility of the methods introduced in these articles, and to compare these methods with earlier methods used by others, I reanalyzed data that had been analyzed earlier by others, and I used the variables as they were used by those with whom I was comparing my work. In one of the many examples, I reanalyzed data that had been analyzed earlier by Coleman (1964) and Theil (1970). There were four variables under consideration in that example; three of them were dichotomous and Coleman (1964) dichotomized the fourth. Theil (1970) used the variables as they were used by Coleman in order to compare his method with Coleman's, and I used the variables as they were used by Coleman and Theil in order to compare my methods with both of theirs. (Reference was also made in my article to Zeisel [1968], who presented the fourth variable in trichotomous form, and to the original source of the data in Stouffer et al. [1949].) A somewhat similar kind of remark applies also to the example in which I reanalyzed data that had been analyzed earlier by Davis (1971). In the two illustrative applications referred to above (namely, in the two examples using the data analyzed by Coleman and Theil and by Davis), and in all the other examples in my articles, I used the variables as they were used earlier by those with whom I was comparing my work.

[17] To facilitate the exposition of the illustrative applications presented in this series of articles, almost all of the variables used in these applications were dichotomous; but almost

RECURSIVE SYSTEMS AND NONRECURSIVE SYSTEMS

As noted earlier, the path diagrams in figures 1 and 2 above provide examples of recursive systems; and most of the discussion in the earlier sections in the present note pertained to such systems. Next let us consider briefly nonrecursive systems and their relationship to recursive systems.

In path diagrams of the kind displayed in figures 1 and 2, the level of variables *A* and *B* might affect the odds pertaining to variable *C*, and the level of variables *A*, *B*, and *C* might affect the odds pertaining to *D*; but the level of *D* cannot affect the odds pertaining to any of the other variables. Consider next the situation where the level of variables *A*, *B*, and *C* might affect the odds pertaining to *D*; and the level of variables *A*, *B*, and *D* might affect the odds pertaining to *C*. In this situation, we have both an equation for the odds $\Omega_{ijk}^{ABC\bar{D}}$ pertaining to *D* (see, e.g., eqq. [5] and [6]) and a similar kind of equation for the odds $\Omega_{ijl}^{ABD\bar{C}}$ pertaining to *C*. Systems of equations of this general kind can be analyzed using related log-linear models, and the methods for doing so have been described in Goodman (1973*a*, 1973*b*). For the sake of brevity, we shall not consider here these systems of equations, but only the related log-linear models.

In order to facilitate comparison with our earlier path diagrams, let us consider two log-linear models: model M_1 which includes the two-factor effects (*AB*), (*AC*), (*BC*), (*BD*), (*CD*) in the four-way table {*ABCD*}; and model M_2 which includes the two-factor effects (*AB*), (*AC*), (*BD*), (*CD*) in the four-way table. When these models are applied to table 2,[18] we obtain likelihood-ratio χ^2 values of 1.46 with 6 degrees of freedom and 6.74 with 7 degrees of freedom for models M_1 and M_2, respectively. Comparing the χ^2 values for models M_1 and M_2, we find that M_1 is more congruent with the observed data than is M_2, and the difference between the models is statistically significant.

We next compare the χ^2 value for model M_1 with the corresponding value for the causal model in figure 2: namely, 1.61 with 7 degrees of freedom. The difference between these two χ^2 values is not statistically significant. The causal model in figure 2 is more parsimonious than is model M_1, since there

all of these articles also included appropriate formulae for the more general case where the variables may be polytomous. The reader is referred to Goodman (1971) for an illustrative application in which one of the variables is trichotomous (with the three categories ordered from soft to hard), and also for the presentation of more general methods that are appropriate when each variable may have any given number of categories, with the categories of each variable either ordered in some specified way or left unordered. Of course, the fact that the methods presented in this series of articles can be applied when each variable can have any given number of categories does *not* at all imply that the results obtained with these methods will be unaffected by the number of categories used with each variable.

18 For the sake of brevity, only table 2 will be considered here. The interested reader can apply these models to table 1 as well (see, e.g., Goodman 1973*a*).

are 7 degrees of freedom pertaining to the causal model in contrast to 6 degrees of freedom pertaining to model M_1.

From the results presented in the two preceding paragraphs, we see that, in the comparison of models M_1 and M_2, because model M_2 did not include the two-factor effect (*BC*) in the four-way table {*ABCD*} and model M_1 did include this effect, there was a statistically significant difference between M_1 and M_2; whereas, in the comparison of model M_1 and the causal model in figure 2, although the causal model did not include a corresponding two-factor effect between variables *B* and *C* in the three-way table {*ABC*}, the difference between these two models was not statistically significant. Under the assumed order of priority of the variables in the causal model in figure 2, this causal model will be preferable to model M_1,[19] since the former model is more parsimonious.

THE MAGNITUDE OF THE EFFECTS

We shall now briefly reflect upon the magnitude of the effects among the variables in different models. Thereby, we shall gain some insight into, say, the reason why the two-factor effect between variables *B* and *C* in the three-way table {*ABC*} could be deleted, as was done in figure 2, whereas a corresponding two-factor effect in the four-way table {*ABCD*} could not be deleted.[20]

Figure 3 displays the path diagram for the log-linear model M_1 applied to the four-way table (table 2), in the case in which each of the four variables in the table can be affected by the other variables in the table.[21] To

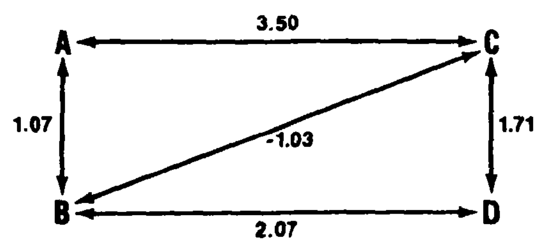


FIG. 3.—Path diagram for the log-linear model M_1 applied to table 2

[19] Under the assumed order of priority of the variables, model M_1 is equivalent to a causal model of the kind displayed in fig. 1 herein. (For further details, see Goodman 1973*a*.)

[20] As noted in the preceding section, with the deletion of the two-factor effect (*BC*) in the four-way table {*ABCD*}, we obtained model M_2 which was not as congruent with the observed data as was model M_1. And the difference between the models was statistically significant.

[21] For the corresponding path diagram obtained when variables *A* and *B* are viewed as being prior to variables *C* and *D*, see Goodman (1973*a*). For expository purposes, we present in fig. 3 here only the case mentioned in the text.

compare the magnitude of the corresponding effects presented in figures 2 and 3, we can apply some results introduced in Goodman (1972*b*) which show in which direction the magnitude of an effect will change when variables are deleted (i.e., collapsed) or inserted.

Before comparing the magnitudes of the corresponding effects in figures 2 and 3, let us first review the results introduced in Goodman (1972*b*). The following kind of sign-rule was presented there: When there is no three-factor interaction in a three-way $2 \times 2 \times 2$ table $\{ABC\}$, then the two-factor effect (AB) in the two-way marginal table $\{AB\}$ will be larger than, smaller than, or equal to a corresponding effect in the three-way table, if the product of the two-factor effects (AC) and (BC) in the three-way table is positive, negative, or zero, respectively. Table 4 displays the possible consequences derived with this sign-rule. From table 4 we see that, for example, when (AC) and (BC) have the same sign, if the (AB) effect in table $\{ABC\}$ is negative, then this effect will increase when variable C is collapsed so that (AB) will go from being negative to being either less negative or zero or positive. Similarly, we see that, for example, when (AC) and (BC) have different signs, if the (AB) effect in table $\{ABC\}$ is positive, then this effect will decrease when variable C is collapsed so that (AB) will go from being positive to being either less positive or zero or negative. We can now apply here the sign-rule described above and various generalizations of this rule introduced in Goodman (1972*b*).

With the comparison of χ^2 values for models M_1 and M_2 in the preceding section, we found that an effect of -1.03 pertaining to the two-factor effect (BC) in the four-way table $\{ABCD\}$ made a statistically significant contribution, and so we preferred model M_1 (which included this effect) to

TABLE 4

THE SIGN-RULE APPLIED TO THE TWO-FACTOR INTERACTION (AB) IN THE THREE-WAY $2 \times 2 \times 2$ TABLE $\{ABC\}$, WHEN VARIABLE C IS COLLAPSED

SIGN OF THE INTERACTIONS IN TABLE $\{ABC\}$		CHANGE IN (AB) WHEN VARIABLE C IS COLLAPSED
(AC)	(BC)	
+	+	Increase
−	−	Increase
+	−	Decrease
−	+	Decrease
0	+ or −	No change
+ or −	0	No change
0	0	No change

model M_2 (which did not include the effect). On the other hand, the −1.03 effect in figure 3 could be replaced by a zero effect when the three-way table {*ABC*} was analyzed as part of the overall causal model in figure 2. Applying a generalization of the sign-rule in Goodman (1972*b*), with the deletion of variable *D*, the two-factor effect (*BC*) will increase (from −1.03 to approximately zero) since the product of the (*BD*) effect and the (*CD*) effect is positive.[22]

Note that the (*BD*) and (*CD*) effects in figure 3 are equal in magnitude to the corresponding effects in figure 2. This is a consequence of the fact that both the log-linear model M_1 and the overall causal model in figure 2 imply the same equation (6) for the odds $\Omega_{ijk}^{ABC\bar{D}}$.[23]

Note also that the (*AB*) and (*AC*) effects in figure 3 are somewhat smaller than the corresponding effects in figure 2. This is a consequence of the following facts: With the deletion of variable *D*, the (*AB*) and (*AC*) effects in the three-way table {*ABC*} will be equal in magnitude to the corresponding effects in the four-way table {*ABCD*} since the (*AD*) effect was zero in figure 3;[24] and as noted earlier in this section, the (*BC*) effect will be increased from −1.03 to approximately zero. Actually, the (*BC*) effect in the three-way table {*ABC*} turns out to be close to zero but slightly positive; and when this effect is set equal to zero (as it is in the analysis of table {*ABC*} under the overall causal model in figure 2), then the only two-factor effects in table {*ABC*} are the (*AB*) and (*AC*) effects, each of which can then be measured (under the causal model) using the corresponding two-way marginal tables {*AB*} and {*AC*}, respectively. The (*AB*) effect in the two-way table {*AB*} will be larger than a corresponding effect in the three-way table because the product of the (*AC*) effect and the (*BC*) effect is positive; and the (*AC*) effect in the two-way table {*AC*} will be larger than a corresponding effect in the three-way table because the product of the (*AB*) effect and the (*BC*) effect is positive.

[22] The negative (*BC*) effect (namely, −1.03) in the four-way table {*ABCD*} analyzed in fig. 3 can be used to explain why a corresponding effect in the three-way marginal table {*ABC*} is approximately zero; or, conversely, the zero effect between variables *B* and *C* in the three-way table {*ABC*}, analyzed as part of the overall causal model in fig. 2, can be used to explain why there is a negative (*BC*) effect in the four-way table {*ABCD*}. The latter form of explanation is preferable since the overall causal model in fig. 2 is more parsimonious than is model M_1 in fig. 3.

[23] Model M_1 also implies that there is no three-factor interaction (*ABC*) in the three-way table {*ABC*}, and the overall causal model in fig. 2 implies that variables *B* and *C* are conditionally independent of each other, given variable *A* (see Goodman 1973*a*).

[24] Using a generalization of the sign-rule referred to earlier, the (*AB*) effect will remain unchanged because the product of the (*AD*) effect and the (*BD*) effect will be zero; and the (*AC*) effect will remain unchanged because the product of the (*AD*) effect and the (*CD*) effect will be zero.

APPENDIX

The purpose of this Appendix is to clarify why the structure of a system of quantitative variables will usually be very different from the structure of the corresponding dichotomized variables. To facilitate this exposition, we shall consider the case where the structure of the system of quantitative variables can be described in terms of the concept of conditional independence. Under conditions specified below, if conditional independence holds true for specified quantitative variables in the system, then conditional independence will *not* hold true for the corresponding dichotomized variables.

Let X, Y, and Z denote three quantitative variables, and let X', Y', and Z' denote the corresponding dichotomized (or polytomized) variables, with X, Y, and Z different from X', Y', and Z', respectively. Suppose that X and Y are conditionally independent of each other, given variable Z. For simplicity, we shall consider the case in which $X = a + bZ + \epsilon_1$ and $Y = c + dZ + \epsilon_2$, with $b \neq 0$, $d \neq 0$, and the usual error terms ϵ_1 and ϵ_2 are independent of each other. In this case, with X and Y conditionally independent of each other, given variable Z, we shall find that X and Y will *not* be conditionally independent of each other, given variable Z', *and* X' and Y' will *not* be conditionally independent of each other, given variable Z'; but X' and Y' will be conditionally independent of each other, given variable Z.

To understand why X and Y are *not* conditionally independent of each other, given variable Z', consider first a specified value of Z' and each of the possible values of Z that can yield this specified value of Z'. For the specified value of Z', the conditional joint distribution of X and Y (given this specified value of Z') can be obtained by considering the conditional joint distribution of X and Y obtained for each possible value of Z that can yield the specified value of Z', and then weighting these joint distributions in accordance with the distribution of the corresponding possible values of Z. Thus, in the special case in which, say, two possible values of Z can yield the specified value of Z', then the conditional joint distribution of X and Y (given this specified value of Z') is a mixture of two conditional joint distributions that have different expected values for X (since $b \neq 0$) and different expected values for Y (since $d \neq 0$); and in the mixture thus obtained variables X and Y will be positively correlated when $bd > 0$ and negatively correlated when $bd < 0$.

As noted above, variables X and Y are *not* conditionally independent of each other, given variable Z', when $b \neq 0$ and $d \neq 0$. However, in the special case in which $b = 0$ and/or $d = 0$, then X and Y are conditionally independent of each other, given variable Z'.[25]

[25] To understand why this is so, return to the discussion presented in the preceding paragraph, and note now that the two conditional joint distributions in the mixture described

The comments in the two preceding paragraphs pertain to the analysis of the conditional joint distribution of X and Y, given a specified value of Z'. These comments can also be extended directly to the analysis of the conditional joint distribution of X' and Y', given a specified value of Z'.

Consider next the analysis of the conditional joint distribution of X' and Y', given variable Z. It is easy to see that if two quantitative variables are independent of each other, the corresponding dichotomized (or polytomized) variables will also be independent of each other. (Consider the joint distribution of the two quantitative variables and the corresponding joint distribution obtained from it when the two variables are dichotomized or polytomized.) Thus, when X and Y are conditionally independent of each other, given variable Z, then X' and Y' will also be conditionally independent of each other, given variable Z.

The results presented in this Appendix could have been stated in more general terms, but we shall not go into these matters here.

REFERENCES

Bishop, Y. M. M. 1969. "Full Contingency Tables, Logits, and Split Contingency Tables." *Biometrics* 25 (June): 383–400.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: M.I.T. Press.

Boudon, R. 1968. "A New Look at Correlation Analysis." Pp. 199–235 in *Methodology in Social Research*, edited by H. M. Blalock, Jr., and A. Blalock. New York: McGraw-Hill.

Campbell, D. T. 1963. "From Description to Experimentation: Interpreting Trends as Quasi-Experiments." Pp. 212–42 in *Problems in Measuring Change*, edited by C. W. Harris. Madison: University of Wisconsin Press.

Coleman, J. S. 1964. *Introduction to Mathematical Sociology*. New York: Free Press.

Davis, J. A. 1971. *Elementary Survey Analysis*. Englewood Cliffs, N.J.: Prentice-Hall.

Duncan, O. D. 1975. "Partitioning Polytomous Variables in Multiway Contingency Analysis." *Social Science Research* 4 (December): 167–82.

Gillespie, M. W. 1978. "The Application of Log-linear Techniques to Recursive Models: Comment on Reynolds." *American Journal of Sociology* 84 (November): 718–22.

Goodman, L. A. 1970. "The Multivariate Analysis of Qualitative Data: Interactions among Multiple Classifications" *Journal of the American Statistical Association* 63 (March): 1091–1131.

———. 1971. "The Analysis of Multidimensional Contingency Tables: Stepwise Procedures and Direct Estimation Methods for Building Models for Multiple Classifications." *Technometrics* 13 (February): 33–61.

———. 1972*a*. "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables." *American Sociological Review* 37 (February): 28–46.

———. 1972*b*. "A General Model for the Analysis of Surveys." *American Journal of Sociology* 77 (May): 1035–86.

———. 1973*a*. "Causal Analysis of Data from Panel Studies and Other Kinds of Surveys." *American Journal of Sociology* 78 (March): 173–229.

———. 1973*b*. "The Analysis of Multidimensional Contingency Tables When Some Variables Are Posterior to Others: A Modified Path Analysis Approach." *Biometrika* 60, pt. 1: 179–92.

---

there will have the same expected values for X when $b = 0$ and the same expected values for Y when $d = 0$; and in the mixture thus obtained variables X and Y will be independent of each other (when $b = 0$ and/or $d = 0$).

———. 1974. "The Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable. Part I—a Modified Latent Structure Approach." *American Journal of Sociology* 79 (March): 1179–1259.

———. 1975. "The Relationship between Modified and Usual Multiple Regression Approaches to the Analysis of Dichotomous Variables." Pp. 83–110 in *Sociological Methodology 1976*, edited by D. R. Heise. San Francisco: Jossey-Bass.

——— 1978. *Analyzing Qualitative/Categorical Data: Log-linear Models and Latent-Structure Analysis*. Cambridge: Abt.

Lazarsfeld, P. F. 1948. "The Use of Panels in Social Research." *Proceedings of the American Philosophical Society* 92 (November): 405–10

———. 1970. "A Memoir in Honor of Professor Wold." Pp 78–103 in *Scientists at Work*, edited by T. Dalenius, G. Karlsson, and S. Malmquist Uppsala: Almquist & Wiksells.

Lazarsfeld, P. F., B. Berelson, and H. Gaudet. 1968. *The People's Choice*. 3d ed. New York: Columbia University Press.

Lipset, S. M , P. F. Lazarsfeld, A. H. Barton, and J. Linz 1954. "The Psychology of Voting: An Analysis of Political Behavior." Pp. 1124–75 in *Handbook of Social Psychology*, edited by G. Lindzey. Vol. 2. Reading, Mass · Addison-Wesley.

Reynolds, H. T 1977. "Some Comments on the Causal Analysis of Surveys with Log-linear Models." *American Journal of Sociology* 83 (July): 127–43.

Stouffer, S. A., E. A. Suchman, L. C. Devinney, S A. Star, and R M. Williams, Jr. 1949. *The American Soldier: Adjustment during Army Life. Studies in Social Psychology in World War II*. Vol. 1. Princeton, N.J., Princeton University Press.

Theil, H. 1970. "On the Estimation of Relationships Involving Qualitative Variables." *American Journal of Sociology* 76 (July): 103–54

Zeisel, H. 1968. *Say It with Figures*. 5th ed. New York: Harper & Row.

Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions[1]

David Snyder
Indiana University

Edward L. Kick
University of Utah

This paper addresses world-system/dependency theories of differential economic growth among nations. We grant that such perspectives have considerable analytic potential but have serious reservations concerning their current empirical status. Our criticisms focus particularly on the absence of evidence on the theoretically specified structural positions (core, semiperiphery, periphery) in the world system and the dynamic relations among them. After indicating why extant quantitative studies that claim to represent "position" are inadequate, we propose that blockmodel analyses of social structure through multiple networks address world-system formulations far more appropriately. We present a blockmodel of the world system circa 1965 that is based on four types of international networks: trade flows, military interventions, diplomatic relations, and conjoint treaty memberships. While we invite replications with additional network data, this blockmodel provides strong evidence for a core-semiperiphery-periphery structure. We then report regression analyses of the effects of these structural positions on nations' economic growth (change in GNP per capita) from 1955 to 1970. Net of other plausible determinants, these effects are large in magnitude and entirely consistent with world-system/dependency theories. Further analyses reinforce the interpretation of these findings as the structural, accumulative advantage of location in the core over that in the periphery. Substantively, our results suggest that exogenetic theories of economic growth are even more powerful than previous analyses have indicated. Moreover, they demonstrate the natural wedding of a conceptual framework (the world system) with an empirically grounded theory of social structure (blockmodel analysis), which has applicability much beyond issues of economic growth.

INTRODUCTION

Differential economic growth among countries constitutes both a central dimension of international stratification and an important determinant of other characteristics (transnational power, individuals' life chances) that

[1] Financial support for this research was provided by the Center for International Policy Studies, Indiana University, under an institutional grant from the Ford Foundation We

are delimited by geopolitical boundaries (e.g., Horowitz 1972). The arena of economic growth has now become a widely used testing ground for alternative theoretical approaches that also frequently address broader issues concerning processes of modernization, societal development, etc. (Portes 1976). Recent perspectives emphasizing the "world economy" as the appropriate level of inquiry and the dynamic, exploitative linkages among its constituent political units as a substantive corollary appear to have considerable analytic power (see Wallerstein 1974; Chirot 1977; Portes 1976; Chase-Dunn 1975; Rubinson 1976). These world-system/dependency theories therefore serve as a sharp counterpoint to the ontogenetic assumptions of more conventional developmental treatments (Rostow 1960; Inkeles and Smith 1974; see Portes [1976] and Chase-Dunn [1975] for cogent reviews of both literatures).[2]

While several analyses attempt to adjudicate hypotheses derived from the alternative theoretical models of economic growth (e.g., Chase-Dunn 1975; Delacroix and Ragin 1978), that is not our purpose here. In fact, we treat developmental theories largely as straw men. That is not because they lack any merits. Available arguments and evidence (Portes 1976; Delacroix and Ragin 1978) suggest that national development and economic growth in particular reflect a complex interaction of inter- *and* intranational factors (especially state strategies and educational institutions). Nevertheless, we are in accord with Portes's (1976, p. 80) assessment that "newer perspectives [world-system/dependency approaches] . . . represent a potentially more useful guide for future investigation." An exogenetic vantage point can easily incorporate domestic factors (world-system approaches consider this to be mandatory), while the converse would be conceptually more difficult.

Consequently, we take world-system/dependency analysis as our point of substantive departure in this paper. But the main premise underlying our treatment is that, despite its analytic potential, the empirical status of this

---

are also grateful to Mark Hayward and Pamela Hecht for software development and consulting, to Steven Brams for making available his diplomatic exchange data, and to Jacques Delacroix for providing some of the cross-national measures used here For helpful comments (not all of which we were able or willing to incorporate) we wish to thank Howard Aldrich, Christopher Chase-Dunn, Jacques Delacroix, Lawrence Hazelrigg, Michael Hout, Paula Hudis, Larry Isaac, Peter Oberto, Harvey Starr, and the anonymous referees. The authors are responsible for all views expressed here.

[2] We are well aware that longstanding theories of imperialism (e g , Lenin's) underlie world-system/dependency analyses (see Portes 1976; Chase-Dunn 1975; Chirot 1977) However, recent variants of such arguments (Wallerstein 1974; Chirot 1977) are more directly central to our current aims. We also recognize that world-system and dependency theories are not necessarily identical. The latter arguably refer more strictly to certain structural aspects of Latin American nations specified by a group of radical economists from those areas (see McDaniel [1976–77] for a review) These distinctions are again unimportant for present purposes, and other treatments (e.g., Delacroix 1977) consider world-system/dependency analysis as a single line of argument.

formulation remains open to serious challenge. We shall develop in detail criticisms concerning the absence of clear operational criteria for identifying the location of nations in the world system, and the corresponding lack of evidence that exploitative structural relations in that system operate as theoretically specified. We also discuss related problems which characterize recent quantitative studies of the effects of dependency on economic growth in the post–World War II period. We argue that, despite explicit claims to the contrary, these studies do not adequately measure the concepts of "position" and "control" in the world system and do not fully test the relevant hypotheses. We propose instead that the blockmodel analyses developed by White and his associates (e.g., White, Boorman, and Breiger 1976) more closely capture these structural aspects of the world economy. The bulk of the paper reports two separate but related investigations that bear on world-system theory. The first is a multiple-network (blockmodel) treatment of transnational flows in the 1960s. This analysis delineates structure in the world system and various positions within it. The second addresses the specific issue of the effect of structural position in the system on nations' economic growth from 1955 to 1970.[3] Although these analyses are empirically independent, they converge on similar substantive conclusions concerning the operation of the world economy in the postwar period. We focus primarily on economic growth but also suggest the utility of our line of inquiry for other hypotheses suggested by exogenetic perspectives on domestic processes.

EXPLAINING DIFFERENTIAL ECONOMIC GROWTH AMONG NATIONS

Several cogent expositions of the alternative theoretical statements have recently appeared (e.g., Chase-Dunn 1975; Portes 1976; Chirot 1977; Delacroix and Ragin 1978). Therefore, we briefly summarize the relevant substantive themes and then focus more intensively on the empirical work that is of primary interest here.

Many developmental explanations treat economic growth largely as a

[3] In order to avoid terminological confusion at the outset: *Wealth* refers to the average economic welfare of a population (which we operationalize as GNP per capita) at a single point in time. *Economic growth* designates only changes over time in wealth We use the term *economic development* (only in a few necessary instances) in its conventional, loose sense of a complex of institutional changes that typically facilitate sustained economic growth. Economic growth and economic development are not identical; the former can occur without the latter (e.g., the Spanish bullion trade). *Differential economic growth* simply identifies variations among nations on this criterion. *Position* (in its verbal world-system and formal blockmodel usages) and economic growth are separate though plausibly empirically related dimensions of international stratification. Position is a discrete variable whose values (e.g., core) are understood in relation to other values (e.g., periphery). Economic growth, wealth, etc. are continuous variables whose values (e.g., $500 per capita) have independent meaning. Perhaps the best analogue is the distinction between class and income (see Wright and Perrone 1977).

process endogenous to nations and as an ordered sequence in which currently poor countries can repeat the strategies and experiences of wealthy nations. For example, neoclassical economic formulations (e.g., Rostow 1960) stress the importance of internal mechanisms that reduce consumption and increase savings, investment, and capital formation.[4] That is not to claim that such processes operate entirely independent of exogenous influences. Economic development in poor nations is enhanced by importation of investments and by more direct forms of foreign aid (both in money and technical knowledge). What has loosely been termed "modernization theory" in sociology makes roughly similar assumptions, although its substantive focus is more on shifts in attitudes, aspirations, and values engendered by contact with modern institutions (Inkeles and Smith 1974).[5]

Developmental explanations have been criticized on a variety of general theoretical (Bendix 1967; Portes 1976) and specific empirical grounds (e.g., Landes 1969, pp. 77–80).[6] More important in generating alternative formulations, however, is simply the prima facie evidence: Despite decades (in some cases, centuries) of cultural contact, foreign investment, etc., there has been relatively little change in international economic stratification. With few exceptions, poverty is endemic in the Third World while rich nations continue to increment their wealth more rapidly (Horowitz 1972; Portes 1976, pp. 56–59).[7]

World-system/dependency theories (e.g., Wallerstein 1974; Chirot 1977; Galtung 1971) attempt to explain those phenomena and are characterized by the following major themes. First, the world system is the appropriate point of conceptual orientation. The behavior and experiences of its constituent geopolitical units depend fundamentally on features of the system

[4] According to Rostow, an investment rate of roughly 10%–12% of net national product is required for the "takeoff" into "self-sustained" economic growth He also develops a parallel but secondary model of intersectoral linkages, which we ignore here

[5] There are differences between evolutionary theories such as Rostow's and theories emphasizing development of values (see Portes 1976). We combine them here because of similar assumptions concerning economic growth as a largely endogenous process and the positive effects of international contact that does occur We also acknowledge that some theories subsumed within the developmental approach (Weber on religious values and capitalism is probably most widely known) do not make all the assumptions we have attributed to this line of argument. Although our characterization of developmental theories is therefore a restricted one, it pertains to most arguments that inform current analyses of economic growth.

[6] For example, the general criticisms typically involve arguments that structural changes over time (e.g., in the scale of enterprise) render it impossible for poor nations to repeat the historical experience of rich ones. Among a variety of empirical points, Landes cites evidence that early capital formation in Britain and France most likely never approached Rostow's "takeoff" level.

[7] Given strict assumptions of unilinear processes of economic growth, there may be no reason to expect shifts in countries' relative wealth. However, the positive effects of international economic and value transmission predicted by most current developmental approaches imply that some reordering should occur as a function of those factors.

as a whole (e.g., a capitalist world economy) which reflect transnational linkages (Wallerstein 1974). Second, the modern world system is composed of three structural positions: core, semiperiphery, and periphery (see Wallerstein [1974, pp. 301–2, 349–50] and Chirot [1977, p. 13] for defining characteristics of each).[8] Third, these labels are not merely descriptive. They indicate an international division of labor in which the core is linked to the periphery (and semiperiphery) in dynamic and exploitative ways. Chase-Dunn (1975), Delacroix (1977), and Portes (1976) summarize several mechanisms that in principle reinforce structural barriers against economic growth in peripheral countries. For example, net resource flows are lost to the core as returns on investments. Additionally, peripheral economies are poorly integrated because of production that is both externally oriented and specialized in raw-material commodities. Other arguments deal not with economic development per se, but with increased inequality within nations as a function of national elite control of the export economy and alliances with core powers (Chase-Dunn 1975; Rubinson 1976). Finally, this world system is currently a capitalist world economy (Chirot 1977) and has been for several centuries (Wallerstein 1974). While the bases of this conclusion are obvious from the foregoing discussion, it is useful to recognize (as Chirot and Wallerstein do) that alternative forms of the world system are possible.

Empirical evidence concerning the development of the modern world system is fragmented and descriptive because long-run historical data are typically unavailable (Wallerstein 1974, p. 8). However, several efforts address the effects of indicators of international dependency on economic growth in the post–World War II period. More complete data for this recent period permit using rigorous analytic methods to test hypotheses derived from world-system/dependency theory. For example, Chase-Dunn (1975) uses a sample of poor countries and a panel regression model of change in GNP per capita from 1950 to 1970 to examine neoclassical and dependency models of the impact of foreign investment on economic growth and inequality. His most important finding is that the influence of domestic investment (capital formation) is positive and that of foreign investment negative. This result supports the dependency argument that it is not capital per se but the institutional locus of capital that is central. Chase-Dunn also finds expected effects of dependency measures on inequality within nations, as does Rubinson (1976) in a related but more elaborate analysis.[9] Although neither study

[8] Chirot (1974, pp. 179–81) maintains that the semiperiphery/periphery distinction may have less validity and utility in the post–World War II period than previously, because almost all noncore nations are now semiperipheral. While we recognize that possibility, we retain the conventional definitions of core, semiperiphery, and periphery pending any changes necessitated by our analysis of structure and position in the world system.

[9] We are not substantively concerned with intranational inequality in this paper. However, these studies are relevant here insofar as their independent variables attempt to capture dimensions of the world economy.

can directly represent the hypothesized intervening mechanisms, they provide clear empirical support for world-system/dependency arguments in the postwar period.

This is not to claim that all empirical results in these studies are consistent with world-system hypotheses. Delacroix (1977) reports that extent of specialization in export of raw materials (a measure that addresses world-system arguments of Galtung and Wallerstein) has no effect, net of initial wealth and secondary school enrollment, on economic growth. Given the poor performance of the world-system variable and the important effects of the endogenous measures, he properly cautions against dismissing ontogenetic processes from models of development. In a similar vein, Delacroix and Ragin (1978) analyze the influence of one form of cultural imperialism (film imports). They find that Westernized information flows have a negative effect on economic growth, but one that can be deflected by endogenous factors (particularly "mobilizing regimes"). Substantively, their results suggest that economic growth of nations is a joint function of exogenetic and ontogenetic variables.

THE EMPIRICAL STATUS OF WORLD-SYSTEM/DEPENDENCY THEORY: A CRITICAL APPRAISAL

We view the studies cited above as extremely important in establishing, though sometimes with qualifications, the merits of world-system/dependency explanations of econmic growth during the past few decades. Nevertheless, we propose three related criticisms concerning the empirical status and quantitative treatments of these arguments, although all points do not necessarily apply to each study.

1. There are no operational criteria for classifying countries according to core, semiperipheral, or peripheral location in the system or for assessing temporal shifts in nations' positions. It is obvious that multiple dimensions underlie the concept of position and that such positions are correlated with but by no means identical to economic development (Chirot 1977; Galtung 1971; Wallerstein 1974). However, we find no precise rules of classification, a problem that is perhaps best highlighted by Chirot's (1977) admitted difficulty in specifying where Communist bloc nations are structurally located in the postwar period. Moreover, there is no strong indication that a three-tiered model (core, semiperiphery, periphery) characterizes the operation of the system as a whole. Put another way, empirical treatments have no clear way in which to validate either the specified number of positions or the structural relations among them. These issues would be less problematic if the three-tiered model were simply a heuristic device. However, at least some world-system arguments indicate that the core-semiperiphery-periph-

ery distinctions are structurally necessary in a capitalist world economy (see Wallerstein 1974, p. 349).

2. Previous regression analyses of dependency hypotheses typically require a choice between two less than optimal research strategies. Some studies limit examination to samples of poor countries (Chase-Dunn 1975; Delacroix and Ragin 1978) because they want to address the effects of economic or cultural imperialism on economic growth. But if core and periphery are dynamically linked in theory, we would argue that empirical treatments should represent not only the "costs" of location in the periphery but also the "gains" of core position. The practical difficulty of doing so is that the substantive meaning of standard dependency measures plausibly differs by location in the world system. As Portes (1976; see also Chirot 1977) points out, countries such as Canada are highly investment dependent but they are also wealthy. The alternative choice—including all nations for which data are available in a single analysis—therefore often (but not always; we would exempt Delacroix [1977] from this criticism) necessitates dubious arguments. For example, Rubinson (1976, p. 654) maintains that greater dollar value of a nation's exports and imports (as a percentage of gross domestic product) "puts the state and its economic actors in a position of less power and control in the world-economy." That argument may hold for poor countries, but we do not find it compelling for nations such as the United States and (especially) Japan, whose core location appears to rest in part on ability to increase their trade in the world market.

3. Despite explicit claims that indicators such as investment dependence, trade concentration, etc. measure "position" or "control" in the world system (Rubinson 1976; Delacroix and Ragin 1978), that is not the case. These are continuous variables that stratify nations on given criteria, and no doubt partially reflect position.[10] But they do not represent such position any more than an individual's income or education measures his or her (discrete) class position (Wright and Perrone 1977; see also White and Breiger 1975, p. 68). Moreover, standard dependency measures do not fully specify the institutional locus of transnational flows. For example, the value of investment dependence of a given (say, peripheral) nation does indicate the degree of foreign investment in that country. However, it does not identify whether transnational capital flows originate entirely, partly, or negligibly from core versus other groups of nations. The differential effect of those origins is theoretically plausible and also promises to have considerable policy relevance. Finally, we want to stress that these arguments do not maintain that the use of continuous dependency indicators in models of the

[10] In making this argument, we follow theoretical proponents of world-system approaches (Wallerstein 1974; Chirot 1977; Galtung 1971) in conceptualizing position as a discrete location in the system and control as one possible form of structural relations among positions or actors.

determinants of economic growth is inappropriate. Trade concentration, investment dependence, etc. do address dimensions of international interaction which are at least partially independent of position as we conceptualize it. There is no necessary reason why these indicators should be less important than position or the origin-destination patterns of transactions in explaining countries' differential economic growth.[11] In fact, it is quite possible that both the magnitudes of international interactions and the positions of actors involved in them must be incorporated into analyses of economic growth.

WORLD-SYSTEM THEORY AND BLOCKMODEL APPROACHES

Taken together, the foregoing criticisms indicate a pressing need for empirical approaches that address world-system theories more closely via identification of structure and position. We have indicated that the substantive themes of the world-system/dependency perspective encompass (*a*) specification of discrete structural positions (core, etc.) in a single system; (*b*) identification of the units in each position and temporal shifts in those locations (e.g., Wallerstein's discussion of Spain's movement from core to semiperiphery); (*c*) structural relations among positions (exploitative links between core and periphery); and (*d*) the consequences of one position versus another for individual nations (e.g., accumulative advantage of core location vs. endemic poverty in the periphery). As White et al. (1976) point out, the general theme that actors' behaviors are influenced by their positions in a social structure is a venerable one and has generated various methods of representing structure and position empirically. White et al. also demonstrate several advantages of their blockmodel treatment over previous approaches to identifying social structure (see also Arabie, Boorman, and Levitt 1978). One reason why we prefer blockmodel analysis relative to alternative procedures is that it constitutes more than simply a technique (though we emphasize its technical aspects and applications here). In providing concrete statements concerning "structure," "position," "role," and relations among these constructs, blockmodel analysis contains the elements for a formal (though still very abstract) theory of social structure.

We consider that the general logic, procedures, and substantive utility of blockmodels are sufficiently well established that they need no extended discussion here (see also Breiger, Boorman, and Arabie 1975; White and

[11] As earlier indicated, there are two separate issues to be addressed (1) a general inquiry into the form and operation of world-system structure in the post-World War II period, in which position is a necessary component; and (2) an analysis of a specific outcome (differential economic growth), in which position is merely expected on theoretical grounds to be an important determinant. Conventional dependency indicators are inappropriate only for the first of these two investigations

Breiger 1975; Boorman and White 1976; Breiger 1976; Mullins et al. 1977). However, three central features warrant emphasis. First, the social structure of a population (arbitrarily defined) is specified by the multiple networks of interaction among members of that population. Second, positions ("blocks") in the structure are aggregates of actors who manifest similar patterns of interaction across all relevant networks. This second feature begins to differentiate blockmodel analysis from conventional (e.g., sociometric) approaches (White et al. 1976, pp. 736–37). Finally, the presence or absence of "bonds" (links on each type of network relation) among positions facilitates inferences concerning the form and operation of the social structure.[12]

These features of blockmodel approaches and the substantive questions posed by world-system/dependency theories are clearly complementary. Such methods plausibly resolve, at least in part, the problems discussed above with respect to empirical treatments of world-system structure. In fact, while blockmodel analysis has been largely applied to networks of individual actors, it has some special advantages concerning international interactions.[13] Such advantages are (1) the relative durability and nonreactivity of recorded information on interactions (e.g., compared with questionnaire data), and (2) the inclusion of almost all (depending on data availability) actors in the population of interest.

CONSTRUCTING A BLOCKMODEL OF THE WORLD SYSTEM: DATA AND SUBSTANTIVE CONSIDERATIONS

Nations are the constituent units or members of the "population" in our analysis. Beyond the practical consideration that network data for geopolitical units are typically available on an international basis, it is also the case that national boundaries sharply delimit the relevant interactions discussed below. Moreover, it is hardly controversial to specify nation states as the most important actors in the modern world system (Wallerstein 1974;

[12] Further aggregation and decomposition of these networks of bonds (analysis of "role structures" [see Boorman and White 1976]) generate higher-order inferences concerning the operation of social systems We do not pursue these additional analyses because they (1) introduce questions concerning roles at a level of complexity not yet substantively addressed by world-system theory, and (2) are especially suited for comparing structures across space and/or time Although we raise possibilities for longitudinal studies of the world system in our conclusions, the network data employed here pertain to system structure at a single time period.

[13] This is not to claim the absence of other network approaches to international flows. The basic substantive ideas that underlie blockmodels of the world system are stated in Haas (1964, p. 53) and Brams (1966, pp. 880–82). And in an innovative analysis Brams (1965, 1966) applies a hierarchical decomposition technique to transnational flows. However, his approach is entirely "sociometric" in its search for "cliques," neglect of isolate nations, and separate analysis of each network. It therefore differs from blockmodel treatments in several important ways (White et al. 1976, pp. 736–37).

Chirot 1977). We operationally define the structure of that system according to four types of international networks: trade flows, military interventions, diplomatic exchanges, and conjoint treaty memberships. We first present the organization, sources, and coding procedures used for each type of network data and then discuss the general substantive considerations and limitations that pertain to our operationalization.

The four networks are represented by separate $N \times N$ matrices of countries. The matrices are all coded in binary form as follows. Each "claim" or "instance of a tie from one member of a population to another" (White et al. 1976, p. 768) is designated "1" in the appropriate cell of the matrix, and the absence of a claim coded "0." For the present analysis, we limited collection of data to a period circa 1965 in order to ensure a large number of cases while retaining the diplomatic relations information (which is available at only one point in time). Our blockmodel analyses are based on 118 nations for which data were available for all networks.[14] Where possible, our coding procedures also utilized a short (and arbitrarily selected) series of years circa 1965 in an attempt to minimize idiosyncratic characteristics of a single year.[15] Details concerning construction of each matrix follow.

Trade.—Each country's yearly dollar value of exports to every other country is reported in a serial publication of the International Monetary Fund and the International Bank for Reconstruction and Development (n.d.). Import data are also published, though they virtually mirror the export figures (there are minor accounting variations). Our criterion for the presence of a "tie" or "claim" is a nonzero value of exports from country A to country B in at least two of the five years from 1963 to 1967 inclusive.[16] We also note that the data contain an underreporting bias. The figures for all "Soviet areas" (Albania, Bulgaria, Cuba, China, East Germany, Czechoslovakia, Hungary, Mongolian Republic, North Korea, North Vietnam, Poland, Rumania, and the Soviet Union) and 21 other nations in the sample are at least partially reconstructed from export and import values of reporting countries. We coded trade ties as present among Soviet area countries,

[14] There is one minor exception: the trade data only were published for Belgium and Luxembourg combined. We attributed the reported flows to each country, which probably slightly overestimates the breadth of their trade networks However, we considered this procedure preferable to dropping both countries from the analysis.

[15] Single-year information would be particularly problematic with the interventions data. For example, given extreme hostility between countries A and B, whether A intervenes into B in 1964 or 1965 (as well as lag time for B's possible retaliation) appears highly susceptible to idiosyncratic factors in the short run.

[16] The smallest nonzero value reported is $100,000. We use the criterion of at least two years to eliminate one-time-only transactions, though very few such instances are actually found in the data. We originally coded the data in several categories, but the vast bulk of trade flows falls into the two extreme categories: no link whatsoever or else large amounts of exports in each year Therefore, we suspect that the actual cutting point used makes little empirical difference.

which appeared more valid than assuming the absence of all such ties. Of the remaining 21, only 11 nations had reconstructed data for more than three of the five years on which our coding is based. Therefore, links among these 11 nations and between any of them and the 13 Soviet areas could not have been reported in the data source. We doubt that this bias is drastic because most of the 11 countries with entirely reconstructed data had very few trade links even with reporting nations.

Interventions.—The Taylor-Hudson (1972) World Handbook II Project coded every military intervention (land, air, or sea incursion of one nation into the territory of another) reported in the *New York Times Index* from 1948 through 1967. We obtained a copy of the individual intervention events data file (described in Inter-University Consortium for Political Research [1971, sec. IV]). Our criterion for a tie is at least one intervention of country A into country B during any of the years from 1960 through 1967 inclusive. We suspect that events of the magnitude of military interventions are nearly fully reported (see Snyder and Kelly 1977). Pearson's (1974) attempt to validate portions of these data suggests that is the case, although he finds one intervention omitted in each of the two years he surveys.

Diplomats.—Steven Brams provided an origin-destination matrix of the number of career-level diplomats in either 1963 or 1964 (exact date varies by country). These data were gathered from lists published by host nations, supplemented with questionnaire information, and are elsewhere described in more detail (Brams 1965, 1966). Our criterion for a tie is at least one diplomat sent from country A to the national capital of country B.

Treaties.—Small and Singer (1969, pp. 264–70) list all international treaties in effect during the years 1946 through 1965, with information on dates of inception and termination (if any) and signatory countries.[17] From their tabular data, we coded the presence of a tie as any conjoint partnership in a treaty by countries A and B during any portion of the years 1960 to 1965 inclusive.

While these four networks clearly do not exhaust the forms of international contact, each of them captures a substantively important dimension of transnational interaction. The centrality of economic and military relations (trade and interventions) in almost all approaches to international system structure is obvious (e.g., the work surveyed in Christopherson [1976]). More specifically, theories of imperialism typically emphasize the domination of the world economy by core powers (though they are not always so designated) and the use or threat of superior armed force as the means of ensuring economic domination. We also consider that diplomatic

[17] These treaties are limited to formal military defense agreements of three types: defense pacts, nonaggression pacts, and ententes (see Small and Singer [1969] for definitions of each, as well as coding criteria and original sources). We used all three types (which do not include general agreements of friendship) in coding our treaty matrix.

exchanges constitute a salient form of information flow in the world system. For sending nations, diplomatic missions provide regular and ostensibly reliable information concerning local economic opportunities, political conditions, etc. in the host country.[18] Such missions may also facilitate trade agreements and sometimes serve as a base for attempts to manipulate local conditions. We attach much less significance to the receiving of diplomats. Whatever information is transmitted about the sending country is both geographically removed from the occurrence of events and "filtered" through representatives of that country. As such, we suspect that such information is actually and perceptually less useful to host nations. Finally, we included the treaty data on the rationale that they would address networks of defense commitments or reflect attempts by some nations (e.g., core powers) to legitimate potential military intervention in others (peripheral countries).

We recognize that operationalizing our approach to structure in the world system via these four networks may be controversial. For example, why employ the diplomat and treaty networks, given the clear precedence of economic and military relations in some influential treatments (e.g., Wallerstein 1974) of world-system structure? Why not substitute types of commodities exchanged, which plausibly capture better the core-periphery division of labor, for the trade network? Our decisions on these issues were based on several conceptual and practical considerations. The most important of these concerns the relationship between blockmodel analysis and world-system theory. Blockmodel analysis is a formal network approach that generalizes across populations and particular substantive issues. In principle, investigators would employ all networks of interaction among members of a population in such analyses, particularly since no given network necessarily has any precedence over any other in the blockmodel procedures. In practice, of course, some selection of interactions is required because of resource and data availability constraints. Therefore, investigators attempt to represent "important" dimensions insofar as they (1) provide the most independent information, and (2) facilitate interpretation of the network patterns of interaction among positions in the blockmodel. For example, blockmodel studies typically include networks of both cooperative (positive affect in the case of interpersonal ties) and conflict (antagonistic) interactions.

The world-system perspective is, among other things, a particular theory of the form and operation of international-system structure (see Christopherson [1976] for alternative models). Although each of our four networks could be conceptualized in terms of core-periphery exploitative dynamics (see the discussion above), it is not necessary to do so. In fact, it would be inappropriate if the generalized approach to social structure

[18] Wallerstein (1974, p. 22) cites an interesting medieval parallel to this argument. In addition, the empirical results we present subsequently provide at least a partial validation of our reasoning concerning diplomatic relations.

(blockmodel analysis) were based solely upon networks specified by a variant (which stresses economic relations, such as Wallerstein [1974] and Chirot [1977]) of a single theory of that structure. Many alternative models of international-system structure also emphasize the importance of noneconomic transactions (e.g., Brams 1966; Christopherson 1976). Moreover, even some analyses that view the world system in essentially core-periphery terms (Galtung 1971) identify system structure as a joint function of economic, military, political, and cultural relations.

While this specification of the relationship between blockmodel analysis and world-system theory justifies the inclusion of noneconomic dimensions such as diplomatic exchanges and treaties, it does not constitute a warrant for excluding more detailed economic transaction data (e.g., types of commodities exchanged, investment capital flows). In this case, data availability influenced our strategic decisions. For the mid-1960s period, information on commodities and other types of economic interactions can be constructed only for a substantially attenuated set of nations (relative to the 118 in the present analysis). We could have employed more recent commodity data, which have greatly expanded coverage. However, information on interventions, diplomats, and treaties is not yet available for the 1970s. We have already indicated the undesirability of deleting such networks. Therefore, we chose to represent multiple dimensions of interaction for nearly the universe of nations, even though our network of economic relations would be a crude one. Moreover, we wanted to avoid the possibility that certain results (e.g., a core-periphery structure) might be based only on the economic indicators most likely to produce them.

This is not to claim that our operationalization is optimal. We invite replications with additional or different networks as they become available. We also recognize that different results could be obtained (though that becomes less likely as more networks add successively less independent information). While such issues are far from trivial, they are subsidiary to our main purpose of demonstrating the potential power of blockmodel approaches to the substantive questions posed by world-system theory.

Two further issues require brief mention preliminary to the presentation of our blockmodel findings. First, our empirical analysis necessarily specifies a single world system that encompasses all nations, although it may identify "isolates" as a unique structural position (see White et al. 1976). Some areas may lie entirely outside the world economy, especially historically (see Wallerstein 1974). However, the fact that all nations had ties to other countries in the 1960s suggests that the assumption of a single world economy during the recent period covered by our data is largely warranted. Put another way, if there are a few exceptions (e.g., stateless societies not included in our sample), they should have no material bearing on our analysis. Second, we want to emphasize that our results bear upon the world system

only at a single point in time. They are not intended to address either the historical development of that system (though knowledge of past occurrences can inform our interpretation) or the mechanisms internal to nation states on which we lack direct measures.[19]

STRUCTURE AND POSITION IN THE WORLD SYSTEM: BLOCKMODEL RESULTS AND DISCUSSION

We applied CONCOR (see Breiger et al. 1975; White et al. 1976) to the matrices described above.[20] Figure 1 displays the number of countries in each successive partition of the data, with blocks designated by arbitrary letter "names." Here we deal only with the 10-block partition of the data (those sets of countries labeled A through F′). While the 10-block split is necessarily consistent with four- and six-block partitions, it presents a more refined picture of the overall structure and also generates the greatest ex-

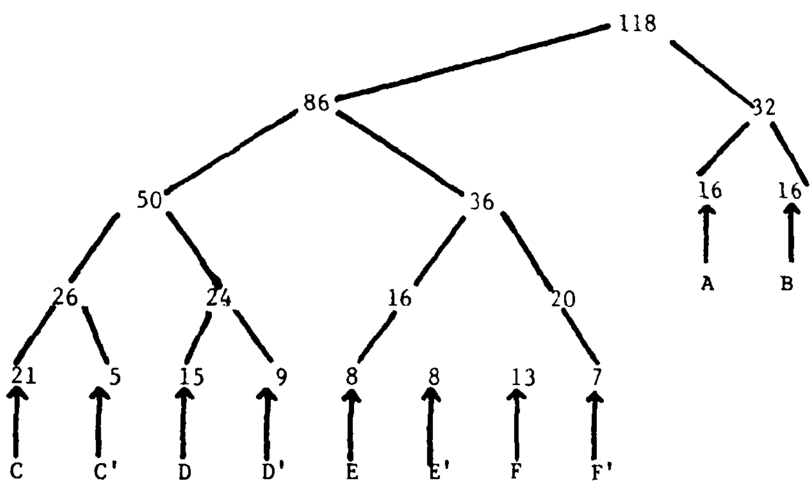


FIG 1.—Blockmodel "tree" of hierarchical decompositions· 118 nations circa 1965. Letters designate block "names" in 10-block model. See table 1 for list of countries located in each block.

[19] By stating these limitations (which also apply to other empirical analyses of world-system issues), we want to avoid metatheoretical criticisms of the form advanced by Bach (1977) and Irwin (1977). We grant the validity of some of their points However, they apply more to analyses of issues (e g , intranational inequality) in which unmeasured domestic attributes such as class structure are central, and which lack time-ordered data. These points are not directly relevant at the level of analysis on which we are operating in this paper

[20] CONCOR involves weaker assumptions than the alternative BLOCKER algorithm discussed by White et al. (1976). We employed CONCOR because this first application of blockmodels to world-system structure is exploratory We also note that studies which employ both algorithms report highly consistent partitions of the data.

planatory power in our subsequent empirical analyses. Table 1 lists the countries in each block for the reader's convenience. However, most of our interpretation does not rely directly on this information.

Table 2 presents the set of binary-image matrices that indicates the presence or absence of "bonds" among blocks on each type of network. The image matrices reflect in more simplified form the patterns of ties among blocks depicted in the density matrices of table 3. With the exception of the treaty matrix (wihch is symmetric by definition), the rows for each dimension indicate the "sending" blocks and the columns the "receiving" ones. For example, the fifth entry in the first row of the trade-density matrix (.744) in table 3 indicates the proportion of actual relative to possible export links from block C nations to block E nations. The binary entries in the image matrices reflect whether each density displayed in table 3 is greater (coded "1") or less (coded "0") than the mean density for that matrix. Although this criterion is of course arbitrary, the overall patterns described below are not highly sensitive to the particular cutting point chosen.

In general, though with some deviations, we interpret the pattern of bonds depicted in the image matrices as a core-semiperiphery-periphery structure

TABLE 1

LISTING OF COUNTRIES IN EACH BLOCK: FOUR-NETWORK BLOCKMODEL (Trade, Interventions, Diplomats, and Treaties) FOR 118 NATIONS CIRCA 1965

Block	Nations
A . .	Chad, Congo (Brazzaville), Congo (Kinshasa), Uganda, Burundi, Rwanda, Somalia, Ethiopia, Malagasy Republic, Morocco, Algeria, Tunisia, Libya, Sudan, United Arab Republic,* Yemen*
B . .	Mali,* Mauritania,* Ghana,* Upper Volta,* Senegal, Dahomey, Niger, Ivory Coast, Republic of Guinea, Liberia, Sierra Leone, Togo, Cameroun, Nigeria, Gabon, Central African Republic
C .	Canada, United States, United Kingdom, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, West Germany, Austria, Italy, Yugoslavia, Greece, Sweden, Norway, Denmark, South Africa, Japan, Australia
C′	Venezuela, Peru, Argentina, Uruguay, South Korea
D . .	Cuba, Ireland, East Germany, Hungary, Cyprus, Bulgaria, Rumania, USSR, Kenya, Iran, Turkey, Iraq, Lebanon, Jordan, Israel
D′ .	Finland, Saudi Arabia, Taiwan, India, Pakistan, Burma, Ceylon, Malaysia, Philippines
E	Panama, Colombia, Ecuador, Brazil, Bolivia, Paraguay, Chile, North Vietnam
E′ .	Haiti, Dominican Republic, Mexico, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica
F. . .	Jamaica, Trinidad and Tobago, Poland, Czechoslovakia, Malta, China (People's Republic), Mongolian Republic, Nepal, Thailand, Cambodia, Laos, New Zealand, Iceland
F′. . . .	Albania, Syria, Kuwait, Afghanistan, North Korea, South Vietnam, Indonesia

* Starred countries in blocks A and B cluster together in further "splits" of the data, although they are not shown as separate blocks in the analyses.

TABLE 2 BINARY "IMAGES" OF TEN-BLOCK MODEL

TRADE

	C	C'	D	D'	E	E'	F	F'	A	B
C	1	1	1	1	1	1	1	1	1	1
C'	1	1	0	1	0	1	0	0	0	0
D	1	0	1	1	0	0	0	0	0	0
D'	1	1	1	1	0	0	0	1	0	0
E	1	0	0	0	1	0	0	0	0	0
E'	1	1	0	0	0	0	0	0	0	0
F	1	0	0	0	0	0	0	0	0	0
F'	1	0	0	0	0	0	0	0	0	0
A	1	0	0	0	0	0	0	0	0	0
B	1	0	0	0	0	0	0	0	0	0

INTERVENTIONS

	C	C'	D	D'	E	E'	F	F'	A	B
C	0	0	1	0	1	0	1	0	1	1
C'	0	0	0	0	1	1	0	1	0	0
D	0	1	1	0	1	1	0	0	1	0
D'	0	0	0	1	0	0	1	0	0	0
E	0	1	1	0	1	0	0	0	0	0
E'	0	0	0	0	1	1	1	0	0	0
F	0	0	0	0	0	0	1	0	0	0
F'	0	0	1	1	0	0	1	0	0	0
A	0	0	0	1	0	0	0	0	1	1
B	0	0	0	0	0	0	0	0	0	0

TABLE 2 (Continued)

DIPLOMATS

	C	C'	D	D'	E	E'	F	F'	A	B
C	1	0	1	1	1	1	1	1	1	1
C'	0	0	0	0	0	0	1	0	0	0
D	0	0	1	1	1	0	1	1	0	0
D'	0	0	1	0	1	0	1	1	0	1
E	0	0	0	0	1	1	1	1	0	0
E'	0	0	0	0	1	1	1	1	1	0
F	0	0	0	0	0	0	1	0	0	0
F'	0	0	0	0	0	0	1	0	0	0
A	0	0	0	0	0	0	1	0	0	0
B	0	0	0	0	0	0	1	0	0	0

TREATIES

	C	C'	D	D'	E	E'	F	F'	A	B
C	1	0	0	0	0	0	0	0	0	0
C'	0	1	0	0	1	1	0	0	0	0
D	0	0	1	0	0	0	0	1	0	0
D'	0	0	0	0	0	0	0	0	0	0
E	0	1	0	0	1	1	0	0	0	0
E'	0	1	0	0	1	1	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
F'	0	0	1	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	1	1
B	0	0	0	0	0	0	0	0	1	1

TABLE 3 TEN-BLOCK DENSITIES

TRADE (Mean Density = .3410)

	C	C'	D	D'	E	E'	F	F'	A	B
C	.943	.895	.908	.889	.744	.714	.630	.612	.622	.500
C'	.829	.560	.320	.422	.250	.525	.215	.114	.087	.100
D	.854	.267	.613	.474	.033	.133	.282	.286	.271	.108
D'	.884	.533	.459	.617	.153	.194	.308	.349	.215	.069
E	.679	.225	.042	.139	.422	.172	.106	0	.023	0
E'	.643	.525	.125	.125	.125	.297	.135	.036	.039	.023
F	.634	.215	.262	.265	.029	.135	.183	.165	.144	.087
F'	.551	.171	.229	.317	.018	.018	.132	.082	.045	.009
A	.604	.037	.225	.181	0	.039	.115	.045	.141	.063
B	.533	.025	.092	.042	0	.008	.038	0	.070	.160

INTERVENTIONS (Mean Density = .0072)

	C	C'	D	D'	E	E'	F	F'	A	B
C	0	0	.016	0	.042	.006	.015	.007	.030	.030
C'	0	0	0	0	.025	.025	0	.029	0	0
D	0	.013	.013	0	.008	.008	.005	0	.017	.004
D'	0	0	0	.025	0	0	.017	0	.007	.007
E	0	.025	.008	0	.047	0	0	0	0	0
E'	0	0	0	0	.031	.016	.019	0	0	0
F	0	0	0	0	0	0	.018	0	.005	0
F'	0	0	.019	.016	0	0	.011	0	0	0
A	0	0	.004	.014	0	0	0	0	.070	.008
B	0	0	0	0	0	0	0	0	0	0

DIPLOMATS (Mean Density = .2644)

	C	C'	D	D'	E	E'	F	F'	A	B
C	.413	.200	.533	.434	.542	.548	.872	.626	.440	.342
C'	.133	0	.067	.044	.200	.225	.400	.257	.138	.075
D	.238	.040	.347	.304	.325	.242	.759	.314	.221	.246
D'	.143	0	.311	.259	.417	.222	.684	.429	.243	.285
E	.173	0	.108	.042	.266	.391	.442	.339	.203	.086
E'	.202	0	.175	.097	.297	.359	.635	.411	.313	.172
F	.117	0	.185	.188	.202	.163	.527	.132	.111	.139
F'	.136	0	.257	.175	.196	.196	.518	.183	.134	.161
A	.119	.050	.154	.118	.133	.133	.447	.143	.129	.133
B	.092	.075	.146	.104	.078	.141	.452	.098	.109	.141

TABLE 3 (Continued)

TREATIES (Mean Density = .1278)

	C	C'	D	D'	E	E'	F	F'	A	B
C	.324	.048	.044	.053	.048	.042	.073	0	.003	0
C'	.048	.480	0	0	.800	.700	0	0	0	0
D	.044	0	.142	.052	0	0	.067	.152	.092	0
D'	.053	0	.052	.025	0	0	.043	.032	.049	0
E	.048	.800	0	0	.875	.875	0	0	0	0
E'	.042	.700	0	0	.875	.656	0	0	0	0
F	.073	0	.067	.043	0	0	.024	.044	0	.010
F'	0	0	.152	.032	0	0	.044	.041	.125	0
A	.003	0	.092	.049	0	0	0	.125	.867	.938
B	0	0	0	0	0	0	.010	0	.938	.938

in which (1) block C constitutes the core; (2) blocks E through B (in the order shown) are the periphery; and (3) block D, and perhaps also C′ and D′, are located in the semiperiphery of the world system.[21] This three-tiered structure is most clearly reflected in the image matrix for trade (see White et al. [1976] for structurally similar patterns with different data). Block C is unambiguously at the core of the world economy. It maintains trade linkages to and from every other block in the system. The density matrix in table 3 similarly indicates that every block has more trade linkages with C than with any other block.[22] At the other extreme, blocks E through B are with minor exceptions integrated into the world economy only through their trade with the core block C. This pattern indicates a peripheral location in the overall structure. Blocks C′, D, and D′ differ from both the "core" and "periphery" (as we designate them). There are mutual trade links among and within these blocks, and these blocks are therefore partially independent of the core block in the world economy. However, blocks C′, D, and D′ are for the most part not directly linked to the peripheral blocks. In addition,

21 We stress that the identification of blocks and inferences concerning system structure are based on a single (simultaneous) partitioning of all four interaction networks. Our discussion below of each network separately is intended only to convey information concerning variations in patterns across types of interaction; it does not imply that each network constitutes a separate blockmodel.

22 CONCOR's dichotomous splits of the raw data (Schwartz 1977) and investigators' decisions of how many partitions to undertake necessarily result in arbitrarily determined numbers of blocks. Therefore, the location of countries in blocks is by no means definitive, and we are not wedded to the precise placement shown in table 1. For example, we could have further partitioned the core block of 21 nations, which would result in a smaller set of "core" nations and a larger set of "semiperipheral" countries. These further refinements do not materially affect our conclusions. For instance, deleting from the core all nations which are arguably semiperipheral would still leave very large differences between the core-periphery and semiperiphery-periphery trade densities depicted in table 3.

they are most "salient" to the core as trade partners (note the discontinuity in the densities after the fourth entry of the first row and col. in table 3). Structurally, blocks C′, D, and D′ reflect a "hangers-on" pattern (White et al. 1976) rather than being part of a simple hierarchy.[23] Substantively, we are therefore inclined to label blocks C′, D, and D′ as the semiperiphery, at least pending examination of the other networks.

The image matrix for interventions is largely consistent with conclusions based upon trade bonds. Although block C does not undertake military interventions into all other blocks, it polices much of the system (note in the density matrix that two of the zero images marginally fail to reach the criterion for a bond). Additionally, block C is the only one that receives no interventions (col. 1). Moreover, the density matrix demonstrates that these zero images are all "true" zeros, and thus entirely independent of the cutting point. No core nations experienced a single territorial incursion during the 1960–67 period covered by these data.[24] Block D is also a ubiquitous intervener, though it differs from C because it is also a frequent target of other blocks (including C). The less global intervention patterns of blocks C′ and D′ reduce our confidence in placing them in a semiperiphery. The remaining blocks vary a good deal in their intervention patterns. In general, there is somewhat more of a hierarchical structure in the intervention than in the trade images. For example, block A countries are intervened into by different, and apparently more powerful, blocks (C and D) than those into which they intervene (D′ and B; block A nations' interventions into each other also manifest the highest density by far in the entire matrix). Similar patterns characterize blocks C′, E′, F, and B. However, indications of any overall hierarchy in the system (except for the military dominance of C) are not strong.

The diplomatic exchange bonds are consistent with our previous hypothesis concerning the differential importance of sending and receiving missions. Block C clearly dominates the sending of these information flows in the system, but it does not receive diplomatic bonds from any other block (col. 1 of the image matrix). Block F is the main receiver, although the meaning of that is ambiguous. Given that block F is also the frequent target of interventions from other blocks, we originally suspected that receiving diplomats might reflect global "crisis locations." But a comparison of the interventions and diplomats matrices indicates no strong support for that contention. As in the interventions matrix, there are modest indications of a hierarchy in

[23] These three blocks do not manifest a clearly dominant pattern of trade over those that lie in the periphery. For example, block B's trade densities are greater among its constituent nations than with any other block except the core.

[24] This particular pattern would not of course be observed during some other periods (e.g., world wars). Multinational wars involving core nations plausibly constitute a primary source of reordering the nations in various positions, if not of altering the positions themselves.

which each location (core, semiperiphery, periphery) initiates but does not receive diplomatic flows from those below it in the status ordering.[25]

In summary, the structure defined by the blockmodel analysis is generally consistent with the core-semiperiphery-periphery model specified by world-system treatments. Block C constitutes the core of the system in each type of network. Only block D is clearly located in the semiperiphery, because it alone reflects the pattern of flows (though to a much weaker extent) that characterizes block C on all interaction dimensions. While blocks C′ and D′ have trade bonds similar to those of block D, their patterns of interventions and diplomatic exchanges are more in line with those of peripheral blocks. Therefore, the placement of blocks C′ and D′ in the semiperiphery is more questionable. Finally, blocks E, E′, F, F′, A, and B clearly comprise the periphery, although there are of course differences among them.

Indeed, there must be differences between all of the blocks because in strict terms the analytic technique has identified 10 separate structural positions (blocks) in the system. Although the exact number of blocks is arbitrary, how can that consideration be reconciled with our claims of a three-tiered model? In our view, the structural relations of blocks E through B are similar enough that they override the observed variations and warrant the "periphery" designation we have attributed to them. However, we also consider that differences among these blocks reflect varying characteristics of different parts of the periphery that may also be substantively important. Such differences plausibly reflect variations in historical experiences, state strategies, etc. Space limitations preclude discussion of these variations here. However, we stress their existence because they suggest that (*a*) national (or international) experiences are not solely determined by world-system position, and (*b*) methods other than those employed here (e.g., case studies, historical analyses) are necessary to explore national variations within particular positions in the system.

THE EFFECTS OF STRUCTURAL POSITION ON THE ECONOMIC GROWTH OF NATIONS, 1955–70

The blockmodel analysis bears only on the form and operation of the world system at a given point in time. Here we integrate the blockmodel findings with our substantive focus on the economic growth of nations. Simply stated, we want to answer two questions: does block location (position) make a difference and, if so, how much of a difference in countries' economic welfare? Therefore, we report multivariate analyses of the effects of block loca-

[25] We ignore the treaty matrix in our discussion because there is not much of substantive importance to be learned from it. Examination of the treaty links among individual nations in the raw-data matrix (not shown here) indicates that the bonds shown in tables 2 and 3 almost entirely reflect regionally based pacts.

tion on economic growth. These analyses constitute strong and independent (of the blockmodel results) tests of world-system/dependency arguments and of our interpretation of the blocks as structural positions.

Following Wallerstein, Chirot, and related arguments, some obvious hypotheses concerning the effects of block location on differential growth may be briefly specified. Block C, net of other factors, should experience the greatest economic growth in the system. By dominating the international trade network, it can largely determine the terms of trade with all other blocks except C′, D, and D′. Given its intervention pattern, block C can maintain its economic interests with the frequent use (and perhaps threat) of military force. Moreover, its apparent invulnerability to such intervention also favors economic growth. And because of its centrality in the diplomatic (information-flow) network, block C should achieve further advantages on the grounds discussed earlier. We expect that block D should experience, again ceteris paribus, the next greatest economic growth. Its trade network is somewhat independent of block C, and D is roughly "second" in initiating interventions and diplomatic missions. Block D should resemble block C in the consequences of those patterns, though to a considerably lesser extent. In addition, since D is a frequent target of outside military intervention, its nations' economic growth may be expected to suffer Finally, the remaining blocks should, according to dependency arguments, manifest the largest decrements as a function of their structural position. Although we expect variations among the peripheral blocks, we are unable to specify their relative magnitudes. For example, blocks A and B are structurally similar on all except the intervention networks. Block A's greater initiation of military incursions, particularly its domination over B, might suggest an economic advantage. But A is also more often the target of interventions, which would suggest economic dislocations. The importance of these differences for economic growth cannot be readily ascertained a priori.

Our empirical analysis employs the change in gross national product (GNP) per capita from 1955 to 1970 as the dependent variable. This interval is long enough to observe substantial shifts and also increments the sample size considerably over an earlier (1950) starting point. Change in GNP/capita is regressed on initial (1955) GNP/capita (to control for its correlation with gain score: see Bohrnstedt [1969]) and on other independent variables of substantive interest. Although a panel or lagged dependent variable model is more typically employed in such analyses (Chase-Dunn 1975; Delacroix 1977), the two forms are equivalent in the parameters of primary interest (Treiman and Terrell 1975, p. 184n).[26] We use the change

[26] Metric partial regression coefficients and standard errors are identical, while the coefficient of the initial GNP/capita variable differs by a constant of 1.0 It should also be noted that the R^2 is generally considerably lower with the change dependent variable, because of the deletion of the autoregression coefficient which usually predominates empirically in the panel model.

score to emphasize our focus on economic growth (see Delacroix and Ragin [1978] for a similar treatment).

We begin with a basic model of change in GNP/capita which follows that of Delacroix (1977). It contains the 1955 GNP/capita and adjusted secondary school enrollment ratio (percentage of the secondary school age population enrolled in secondary education) as explanatory variables. The GNP/capita is the most conventional indicator of average economic welfare.[27] The secondary school ratio more likely influences changes in wealth than primary enrollment over the relatively short period covered by these data. It also minimizes comparability and multicollinearity problems relative to university-level enrollment figures (Delacroix 1977). Equation (1) in table 4 presents estimates of this basic model for the 90 countries for which all data are available. Nations higher in initial wealth experience greater increases over the 15-year period (we assume that this effect reflects a number of unspecified cultural and institutional factors).[28] The impact of secondary enrollment is consistent with Delacroix's findings. Each percentage point increase in 1955 enrollment generates an estimated net $21 gain in GNP/capita.

Equation (2) adds a series of nine dummy variables to the basic model, each representing block location (countries were coded "1" if they were located in the given block and "0" otherwise). Note that one category must be omitted because it is a perfect linear combination of the other dummy variables in the model. We arbitrarily omitted block C. Therefore, the coefficients of location in the other blocks are properly interpreted as the net effects of being in a given block relative to location in the omitted category or "core" block.[29] The results presented in equation (2) are noteworthy in

[27] Data for this portion of the analysis are based on a merged file of information from Taylor and Hudson (1972) and the International Bank for Reconstruction and Development (1971, 1973). The GNP and other measures taken from the latter source are identical to those employed by Chase-Dunn (1975) and Delacroix (1977). Delacroix (1977, p. 800n) provides more detailed information and proper cautions regarding the GNP data. We also recognize that correlated error in the 1955 and 1970 GNP indicators will upwardly bias the slope of initial GNP/capita. But we are not overly concerned with that bias because coefficients of the variables of substantive interest will therefore be conservative estimates.

[28] If dependency arguments are largely correct, then location of course affects initial GNP/capita. Moreover, the persisting influence over time of world-system position is at least partially embedded in the estimated autoregression coefficient. Therefore, assessments of the importance of exogenetic factors suggested by our analyses are almost certainly very conservative. We also want to acknowledge that the longer-run relationship between position and economic growth is reciprocal. Greater or lesser economic growth determines a nation's ability to change or maintain position in the system. However, the substantive questions at issue here and the logic of our research design limit our inquiry to the short-run influence of position on economic growth.

[29] We could of course calculate the effect of the omitted category (minus the weighted average of the coefficients of the included ones) and then estimate deviations from the grand mean rather than from block C. However, that procedure is not important for our current purposes.

TABLE 4

ESTIMATES OF CHANGES IN GNP/CAPITA (1955–70) REGRESSED ON CONTROL VARIABLES AND DUMMY VARIABLES REPRESENTING "STRUCTURAL POSITION" (Location in 10-Block Model)

	PARTIAL REGRESSION COEFFICIENTS (Standard Errors)			
INDEPENDENT VARIABLES[a]	Eq (1)	Eq (2)	Eq (3)	Eq (4)
GNP/capita (1955)	292 (080)**	274 (081)**	266 (082)**	294 (082)**
Secondary school enrollment ratio (1955)	20 99 (2 37)**	13 62 (2 95)**	16 98 (2 76)**	11 18 (3 15)**
Block A		−468 0 (170 4)**		−329 0 (163 1)*
Block B		−507 4 (184 7)**		−449 6 (182 1)*
Block C′		−645 4 (186 0)**		−626 2 (184 1)**
Block D		−422 7 (139 7)**		−386 6 (146 3)**
Block D′		−510 0 (180 1)**		−490 5 (179 9)**
Block E		−542 8 (202 6)**		−504 3 (197 9)*
Block E′		−582 3 (177 5)**		−546 5 (175 9)**
Block F		−310 0 (153 4)*		−257 3 (158 9)
Block F′		−789 5 (204 3)**		−758 3 (204 7)**
Diplomats sent (1963–64)			142 (297)	433 (309)
Interventions into (1960–67)			−1 03 (3 54)	−3 44 (3 59)
Interventions by (1960–67)			−11 47 (4 42)*	−12 68 (4 49)**
Value of exports (1965; millions of U.S $)			036 (018)*	010 (019)
Regression constant	20 33	562 2	51 1	507 6
R^2	68 1%	75 9%	72 7%	79 1%

NOTE —N = 90 nations.
[a] Dependent variable is change in GNP/capita, 1955–70.
* Significant at 05
** Significant at 01.

several respects. The coefficients of the block dummy variables are all in the negative direction expected from a core-periphery model.[80] Moreover, they are enormous: the "cost" of location in a peripheral or semiperipheral block (relative to the core) is roughly $500 per capita over this 15-year span.[81] To be sure, block location is correlated with a country's initial wealth, but these estimates are net of GNP/capita in 1955. It should be stressed that there is no necessary connection between the image and density matrices in the blockmodel analysis and the magnitudes or signs of the effects of the block dummy variables in the regression model. The convergent evidence from these two analyses therefore provides strong support for world-system/dependency theory of system structure and of the most widely investigated consequence of that structure (nations' differential economic growth). We also note that the inclusion of the block variables substantially attenuates the effects of the secondary school enrollment indicator in the "endogenous" model of economic growth. Finally, the series of dummy variables modestly increments the R^2 over the baseline model's explanatory power. This increment is perhaps more compelling in light of Delacroix's (1977, p. 805) conclusion that "in the short run, wealth leads to wealth. It is difficult to show the effect of anything on national wealth."

Despite the strength of the results presented in equation (2), they might still be artifactual. For example, core nations initiate more interventions, diplomatic missions, and trade flows and experience no military incursions. Other countries also vary along these dimensions. Therefore, it is plausible that the effects of block membership and its substantive analogue (structural position in the world system) simply reflect countries' differential values on these dimensions of international interaction. To investigate that possibility, equation (3) temporarily deletes the block dummy variables and adds the total number of interventions initiated and received by each nation (1960–67), diplomats sent (1963 or 1964), and dollar value (in millions) of exports in 1965. We are essentially (but not precisely) including the row and column marginals of the raw data matrices initially used to define the partitioning of nations into blocks. The results indicate that nations which

[80] However, expected differences in the economic costs experienced by nations in semiperipheral versus peripheral locations are not supported by these results While block D (the clearly semiperipheral block) has the second lowest decrement, the general pattern of effects does not sustain that interpretation

[81] Since our data pertain to the quasi universe of nations, standard errors and significance tests are reported solely for reasons of convention and convenience. We are quite willing to rest our conclusions on the magnitudes of the partial regression coefficients. One reader suggested that heteroscedasticity of the disturbances, which is plausible with a skewed dependent variable such as GNP/capita, might in any case bias the reported significance estimates. The regression coefficients that are of primary interest would be unbiased even if the disturbances were not homoscedastic. However, we employed a test for heteroscedasticity developed by Goldfeld and Quandt (1965; see Dutta [1975, pp. 132–36] for details on procedures) and found no evidence of it.

initiate more trade and diplomat flows experience positive economic returns (though small ones for diplomats) on those activities.[32] But the findings for interventions are surprising. The effect of receiving interventions is negligible. Conversely, the impact of interventions by a country is both substantial and negative. Moreover, this effect is "real." Other analyses (not reported in tabular form) indicate that it persists through a variety of estimating equations and is not masking different slopes across blocks (e.g., positive effects of interventions by core nations). Although we are primarily interested in the measures in equation (3) as control variables, this finding warrants brief discussion. We still maintain that nations' bonds in the intervention matrix (e.g., the dominance of block C) contribute to the effects of the block dummy variables in equation (2). However, we believe that the negative effect of the number of interventions by a country in equation (3) could reflect either of two (not necessarily exclusive) phenomena. Interventions may be initiated precisely because the target country is not under firm control, that is, as a response to declines in economic returns extracted from that country. It is also possible that this result indicates a "free rider" problem (e.g., Olson 1965). For example, while the core block may derive collective economic benefits from military imperialism, the particular nations that initiate interventions bear the direct costs of providing or ensuring that collective good.

Finally, equation (4) in table 4 adds the block dummy variables to the controls included in equation (3). Although the effects of block location are slightly attenuated in this expanded model, the absolute and relative magnitudes of the coefficients are similar enough to those of equation (2) that our substantive conclusions need not be materially altered. The effects we attribute to structural position in the system are clearly not due simply to the gross outflow and inflow (i.e., marginal) characteristics of nations' interactions in that system. Therefore, the contention that such effects must be due to the networks of interaction that define the system and position in it becomes even more plausible.

FURTHER ANALYSIS

In order to investigate as exhaustively as possible with available data whether the effects of block location are spurious, we undertook several fur-

[32] Total value of imports is correlated over 99 with exports, so we dropped it from the analysis Therefore, the exports measure more properly reflects total trade volume. We also originally included the numbers of diplomats received and treaty ties for each nation. However, there was no clear substantive rationale for expecting effects on economic growth, and the estimated coefficients were trivial in magnitude. Inclusion of these two variables changes the reported coefficients only in minor ways. We also note that exports and diplomats sent are highly correlated ($r = 88$), and that entering either alone in the equations results in much larger slopes than does including both together.

ther analyses. We first added the block dummy variables to a lagged dependent variable model that included investment dependence per capita (1961–65) and domestic capital formation in 1960 as a percentage of gross domestic product (see Chase-Dunn 1975). We further incorporated all the control variables of equation (4) in table 4, as well as a measure of each country's degree of export concentration in 1965. The latter reflects the concentration of exports (the row entries in our original trade matrix) but not the origin-destination combination captured in the blockmodel partitions. Although the sample size is substantially reduced in this analysis, the effects of block location remain roughly as strong as those reported in table 4. In addition, some treatments (e.g., Chase-Dunn 1975) log transform the dependent variable in order to normalize the skewed distribution of GNP/capita or changes therein. That procedure substantially reduces the fit of the model as a whole (because the autoregression effect diminishes considerably), but the substantive conclusions based on the block-location coefficients are again unaltered by such transformations. We prefer the untransformed GNP/capita estimates (particularly given the absence of heteroscedasticity) because the metrics are so easily interpretable.

In summary, our extensive examination of other important determinants of economic growth identified in recent analyses indicates that the empirical effects of block location are in no way spurious. In our view, substantive interpretation of those effects as the accumulative structural advantage of the core over the periphery (and semiperiphery) of the world system is therefore warranted. As a final test of that interpretation, we estimate equation (2) in table 4 separately for changes in GNP/capita during the periods 1960–70 and 1965–70.[33] Table 5 presents results of those estimates, which may be compared with equation (2) in table 4. The effects of block location remain substantial (particularly for 1960–70). However, the absolute magnitudes of the coefficients of block location decline monotonically as the interval of observation decreases from 15 to five years.[34] These findings suggest that the advantages or costs of particular structural positions operate even in the very short run. But, while they are of modest size over a five-year period, these effects also persist through time and therefore result in

[33] These analyses also address a potential concern with time ordering of the measures. Since the network data and block location variables pertain to the early 1960s, they are observed after some of the economic changes in the 15-year interval (1955–70) have already occurred. However, the findings reported below eliminate that as a plausible explanation for the results presented in table 4.

[34] We also estimated expanded models of the form of equation (4) for changes from 1960 to 1970 and from 1965 to 1970. In these models, the coefficients of the block dummy variables are reduced by amounts roughly proportional to those shown in table 4. In the 1965–70 estimates, the effects for a few blocks (A, B, and F) become almost trivial in magnitude, but the general pattern of results supports conclusions identical to those based on table 4.

TABLE 5

ESTIMATES OF CHANGES IN GNP/CAPITA (1960–70 and 1965–70) REGRESSED ON CONTROL VARIABLES AND DUMMY VARIABLES REPRESENTING "STRUCTURAL POSITION" (Location in 10-Block Model)

	PARTIAL REGRESSION COEFFICIENTS (Standard Errors)	
INDEPENDENT VARIABLES	Change in GNP/Capita (1960–70) Eq (5)	Change in GNP/Capita (1965–70) Eq (6)
Initial GNP/capita (1960 or 1965)	126 (073)	338 (052)**
Secondary school enrollment ratio (1960)	14 90 (2 58)**	7 68 (1 84)**
Block A	−296 5 (143 7)*	−127 8 (96 6)
Block B	−317 7 (149 4)*	−136 8 (99 3)
Block C′ . .	−482 0 (162 1)**	−262 5 (109 1)*
Block D	−425 3 (116 1)**	−210 5 (79 4)**
Block D′	−473 5 (143 2)**	−253 9 (97 1)**
Block E	−380 4 (159 9)*	−201 6 (107 3)
Block E′	−428 7 (153 4)**	−207 1 (103 6)*
Block F	−256 8 (124 0)*	−90 4 (90 1)
Block F′	−684 1 (174.8)**	−361 7 (114 7)**
Regression Constant	351.6	151 1
R^2	75 6%	83 2%

NOTE —N = 105 nations for 1960–70, 106 nations for 1965–70
* Significant at 05.
** Significant at 01

large cumulative differentials in economic growth among nations even over a 15-year period.

CONCLUSIONS AND IMPLICATIONS

Taken together, the analyses reported here provide strong support for world-system/dependency theories of structure in the post–World War II period and of the sources of differential economic growth of nations. The blockmodel results indicate a core-semiperiphery-periphery structure in the world system. Moreover, the effects of structural position on the economic growth of nations from 1955 to 1970 are highly consistent with world-system/dependency formulations, although they do not differentiate between the economic costs of peripheral versus semiperipheral location. Substantively, these findings suggest that exogenetic theories of economic growth may be even more powerful than previous analyses have indicated. This is not to claim that world-system position alone determines the differential economic growth of nations. Both endogenous factors (education) and the magnitudes of some transnational flows (e.g., trade, interventions) have important effects and must also be included in models of economic growth.

More generally, this paper demonstrates a natural wedding of an empirically grounded theoretical approach to social structure (blockmodel

analysis) and a conceptual framework (the world system) that can be fruitfully applied to domestic conflict, inequality, and other substantive issues addressed by world-system treatments (e.g., Chirot 1977). In addition, we have begun designing blockmodel analyses of networks at several points in time which promise to answer questions concerning structural changes in the world system. For example, issues of obvious interest include the structural consequences of the incorporation of many African nations in the system circa 1960 and possible shifts in the location of OPEC countries in the 1970s. We do not mean to imply that such investigations are a substitute for the historical analyses and case studies that are necessary to understand national variations within structural positions and the mechanisms through which system position influences domestic processes. While we have not addressed such lines of inquiry here, our acknowledgment of their importance is more than token. Moreover, data availability will almost certainly limit analyses of the form reported here to the post–World War II period and will therefore preclude examination of long-run historical developments with our procedures. Despite this limitation, the approach employed here should inform research problems concerning the world political economy that we have not yet even considered.

REFERENCES

Arabie, Phipps, Scott A. Boorman, and Paul R. Levitt. 1978. "Constructing Blockmodels: How and Why." *Journal of Mathematical Psychology* 17 (February): 21–63.

Bach, Robert L. 1977. "Methods of Analysis in the Study of the World-Economy." *American Sociological Review* 42 (October): 811–14.

Bendix, Reinhard. 1967. "Tradition and Modernity Reconsidered." *Comparative Studies in Society and History* 9 (April): 292–346.

Bohrnstedt, George W. 1969. "Observations on the Measurement of Change." Pp. 113–33 in *Sociological Methodology 1969*, edited by Edgar F. Borgatta and George W. Bohrnstedt. San Francisco: Jossey-Bass.

Boorman, Scott A., and Harrison C. White. 1976. "Social Structure from Multiple Networks II. Role Structures." *American Journal of Sociology* 81 (May): 1384–1446.

Brams, Steven J. 1965. "Flow and Form in the International System." Ph.D. dissertation, Department of Political Science, Northwestern University.

———. 1966. "Transaction Flows in the International System." *American Political Science Review* 60 (December): 880–98.

Breiger, Ronald L. 1976. "Career Attributes and Network Structure: A Blockmodel Study of a Biomedical Research Specialty." *American Sociological Review* 41 (February): 117–35.

Breiger, Ronald L., Scott A. Boorman, and Phipps Arabie. 1975. "An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling." *Journal of Mathematical Psychology* 12 (August): 328–83.

Chase-Dunn, Christopher. 1975. "The Effects of International Economic Dependence on Development and Inequality: A Cross-national Study." *American Sociological Review* 40 (December): 720–38.

Chirot, Daniel. 1977. *Social Change in the Twentieth Century*. New York: Harcourt Brace Jovanovich.

Christopherson, Jon A. 1976. "Structural Analysis of Transaction Systems: Vertical Fusion or Network Complexity?" *Journal of Conflict Resolution* 20 (December) · 637–62.

Delacroix, Jacques. 1977. "The Export of Raw Materials and Economic Growth: A Cross-national Study." *American Sociological Review* 42 (October): 795–808

Delacroix, Jacques, and Charles Ragin. 1978 "Modernizing Institutions, Mobilization, and Third World Development: A Cross-national Study " *American Journal of Sociology* 84 (July): 123–50.

Dutta, M. 1975. *Econometric Methods.* Cincinnati· South-Western

Galtung, Johan. 1971. "A Structural Theory of Imperialism " *Journal of Peace Research* 8 (2): 81–117.

Goldfeld, S M , and R E Quandt 1965 "Some Tests of Homoscedasticity " *Journal of the American Statistical Association* 60 (September) 539–47

Haas, Ernst B. 1964. *Beyond the Nation-State · Functionalism and International Organization.* Stanford, Calif . Stanford University Press.

Horowitz, Irving L. 1972 *Three Worlds of Development · The Theory and Practice of International Stratification* 2d ed New York Free Press.

Inkeles, Alex, and David H. Smith. 1974 *Becoming Modern · Industrial Change in Six Developing Countries.* Cambridge, Mass Harvard University Press

International Bank for Reconstruction and Development 1971 *World Tables* Washington, D C : IBRD.

———. 1973. *World Economic Atlas* Washington, D C · IBRD

International Monetary Fund and International Bank for Reconstruction and Development. n d. *Direction of Trade, Annual 1963–1967.* Washington, D C International Monetary Fund.

Inter-University Consortium for Political Research 1971 "World Handbook of Political and Social Indicators II· Intervention Data " Mimeographed Ann Arbor, Mich Inter-University Consortium.

Irwin, Patrick H. 1977 "Cross-Sectionalism: Mismatching Theory and Model." *American Sociological Review* 42 (October) · 814–17

Landes, David S. 1969. *The Unbound Prometheus. Technological Change and Industrial Development in Western Europe from 1750 to the Present.* Cambridge· Cambridge University Press.

McDaniel, Tim. 1976–77. "Class and Dependency in Latin America." *Berkeley Journal of Sociology,* vol. 21, pp. 51–88.

Mullins, Nicholas C , Lowell L Hargens, Pamela K. Hecht, and Edward L Kick. 1977 "The Group Structure of Cocitation Clusters A Comparative Study " *American Sociological Review* 42 (August): 552–62

Olson, Mancur 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups.* Cambridge, Mass : Harvard University Press

Pearson, Frederic S. 1974. "Geographic Proximity and Foreign Military Intervention." *Journal of Conflict Resolution* 18 (September) 432–60.

Portes, Alejandro 1976 "On the Sociology of National Development: Theories and Issues." *American Journal of Sociology* 82 (July): 55–85.

Rostow, W. W. 1960. *The Stages of Economic Growth* Cambridge. Cambridge University Press.

Rubinson, Richard. 1976. "The World-Economy and the Distribution of Income within States: A Cross-national Study." *American Sociological Review* 41 (August): 638–59.

Schwartz, Joseph E 1977. "An Examination of CONCOR and Related Methods for Blocking Sociometric Data." Pp. 255–82 in *Sociological Methodology,* edited by David R. Heise. San Francisco: Jossey-Bass.

Small, Melvin, and J. David Singer. 1969. "Formal Alliances, 1816–1965 An Extention of the Basic Data." *Journal of Peace Research* 6 (3): 257–82.

Snyder, David, and William R. Kelly. 1977. "Conflict Intensity, Media Sensitivity and the Validity of Newspaper Data." *American Sociological Review* 42 (February): 105–23.

Taylor, Charles L., and Michael C. Hudson. 1972. *World Handbook of Social and Political Indicators*. New Haven, Conn.: Yale University Press.

Treiman, Donald J., and Kermit Terrell. 1975. "Woman, Work and Wages: Trends in the Female Occupational Structure since 1940." Pp. 157–99 in *Social Indicator Models*, edited by Kenneth C. Land and Seymour Spilerman. New York: Russell Sage.

Wallerstein, Immanuel 1974. *The Modern World-System*. New York: Academic Press.

White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structure From Multiple Networks. I. Blockmodels of Roles and Positions." *American Journal of Sociology* 81 (January): 730–80.

White, Harrison A., and Ronald L. Breiger. 1975. "Pattern across Networks." *Society* 12 (July/August): 68–73.

Wright, Erik Olin, and Luca Perrone. 1977. "Marxist Class Categories and Income Inequality." *American Sociological Review* 42 (February): 32–55.

Some Economic Effects of Revolution: Models, Measurement, and the Cuban Evidence[1]

Michael S. Lewis-Beck
University of Iowa

While theorists have scrutinized the causes of revolution, they have avoided systematic study of its consequences. This paper counters such neglect in a number of ways. Rival models of the effects of revolution are developed. To make testing of these competing models possible, revolution is conceived of as an experimental treatment in an interrupted time-series design. Utilizing this design, economic data on the Cuban Revolution are analyzed. The findings reject three leading models of revolutionary effects—the Conservative, the Marxist, and the Thermidorian. Instead, the results hint that the economic effects of revolution are neither as pervasive nor as disruptive as commonly supposed, at least in the short run.

Revolutionary processes continue to fascinate students of collective behavior. A remarkable amount of attention has been devoted to the general issue of what causes revolution (see, e.g., Brinton 1938; Davies 1962; Gurr 1970; Johnson 1966; Moore 1966; Smelser 1963; Wolf 1969). Curiously, however, theoretical investigations concerning the effects of revolution are scarce. Moreover, the numerous case studies of specific revolutions tend to eschew global models of effects. In fact, Eckstein (1964, p. 28) argues that no questions about revolution "have been more thoroughly neglected by social scientists" than those about consequences. Some 10 years later, Greene (1974, p. 11) echoes this conclusion, observing that "much remains to be done in this area of study, especially from the perspectives of social science." As both critics further note, the few Western scholars who have formally considered revolutionary outcomes have, with the obvious exception of the Marxists, provided largely negative evaluations (Eckstein 1964, p. 2; Greene 1974, p. 11). Through his emphasis on the extensive damage wrought by revolution, Huntington (1968, pp. 308–10) appears a contemporary representative of this tradition.

What explains the comparative lack of research on general models of

[1] Many people have contributed to the development of this paper The comments of anonymous reviewers were especially useful. Warren Dent guided me on certain econometric questions. Jeff Brown, Greg Brunk, and Paul Gough provided invaluable help with data collection and analysis. Finally, Peter Rossi and his colleagues at the Summer Institute in Evaluation Research, University of Massachusetts, first stimulated my interest in interrupted time-series design.

revolutionary effects? Of course, it is difficult to say, but one reason, suggested by Eckstein (1964, p. 6), seems especially important. He argues that revolutions, or "internal wars" in his phrase, are not well suited to the techniques of modern social science, such as social surveys, controlled experiments, and anthropological field work. As he puts it: "Many of them do occur, but nowhere near enough to provide the number of cases one deals with, say, in voting studies. They can hardly be staged experimentally, and tend, for obvious reasons, to deter field workers, interviewers, and pollsters" (1964, p. 6). Revolution does appear to resist analysis, at least through these approaches. Nevertheless, developments in evaluation research suggest a way out of this methodological quandary.

The work of Campbell (1969), especially, has encouraged policy evaluators to regard social reforms as experiments, or quasi experiments. I propose a similar view for revolutions. Such an experimental focus yields a number of benefits. First, consideration of revolution as an experimental treatment immediately fixes attention on outcomes. Second, the particular quasi-experimental approach employed here, that is, an interrupted time-series design, promises a quantitative assessment of revolutionary impacts. Further, this statistical model can be used to test rival theoretical models of the effects of revolution.

In what follows, I begin by formulating three competing models—Conservative, Marxist, and Thermidorian. The models are more formally expressed as possible outcomes in an interrupted time-series design and are tested with Cuban data, which direct attention to the economic effects of revolution. I will look at the revolution's impact on sugar production first and then at economic growth in general.

THEORETICAL AND STATISTICAL MODELS OF REVOLUTIONARY EFFECTS

Because revolutions aim to transform ongoing social processes, they tend to be disruptive events. The theoretical concern here is how revolution disrupts valued social and economic production, in the short run and in the long run. Suppose observations of a nation's economic production fall along a line, as in figure 1, where the x-axis stands for time (T) and the y-axis represents economic production (EP). One observes production as a perfect, positive linear function of time, $EP = f(T)$.

When revolution (R) occurs, the classic hypothesis of impact, for the short run, posits a more or less immediate drop in the level of economic production, as diagramed in figure 2. From this perspective, revolution has an initial cost, indicated by the difference between A' and A. Of course, the usual expectation is that the impact of revolution survives beyond this first jolt. In fact, a common prediction would be that the revolution would introduce structural changes which would alter social and economic output

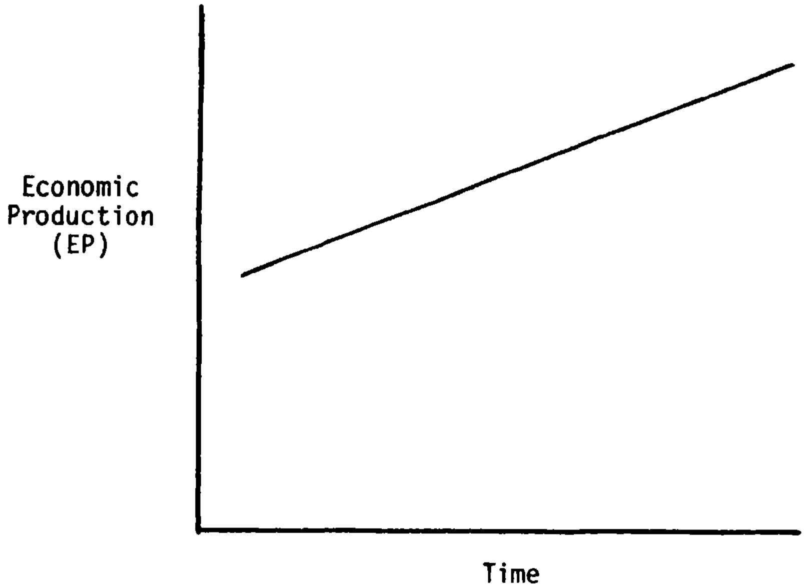


FIG. 1.—No revolution

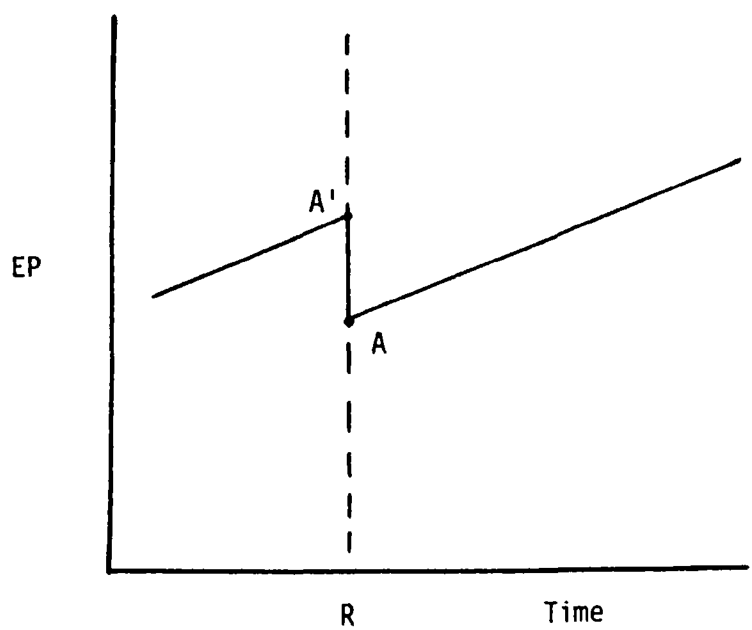


FIG. 2.—Short-run, costly revolutionary effect

in a lasting way. Diagrammatically, these long-run effects appear as a change in the slope (S) of the production trend line following the revolution. Within a linear framework, there are three possibilities: the slope may increase, decrease, or remain the same (see fig. 3). Joining these alternative slope changes ($\Delta_S = +, -, 0$) with the negative change in level ($\Delta_L = -$) yields three models of revolutionary effects. For convenience, I have labeled them Conservative, Marxist, and Thermidorian. I would like to discuss the Conservative model first.

In the Conservative view, revolution is costly, both in the short and long run. This perspective appears to be widely held by Western scholars (Eckstein 1964, p. 2; Greene 1974, p. 11). For example, it would seem congruent with Huntington's (1968, pp. 308–9) assertion that the "economic results of a revolution are almost entirely negative." Beyond the initial disruption of productive forces, revolution inflicts enduring harm on a nation's capacity to produce. The adverse result of the first revolutionary shock is indicated by a drop in economic production: $A' - A$. The diminished slope of the line after the revolution represents the reduced ability of the nation to produce. Revolution, then, becomes a costly endeavor, both at the beginning and indefinitely into the future. Even the prerevolutionary level of output, at A', is not finally restored until point B. Thus, because of these joint short- and long-term effects, the detrimental influence might even extend over generations.

Perhaps the chief theoretical rival to the Conservative model is a Marxist one. According to Marxist theory, revolution creates necessary, substantial short-run costs but yields important long-run benefits. The Marxist revolu-

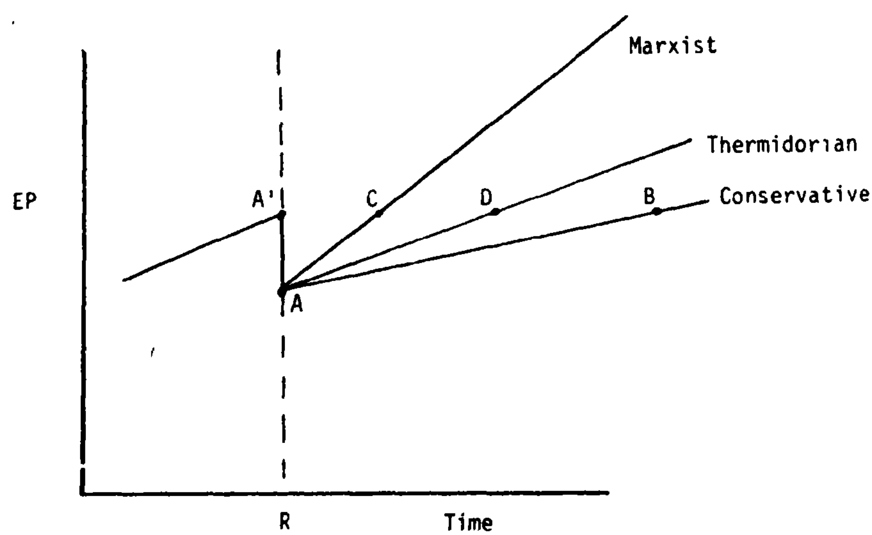


FIG. 3.—Rival models of revolutionary effects

tion, in reordering the ownership of the means of production, generates extreme social and economic dislocation, which immediately hurts production (as with the Russian case). However, the new relations established in the socialist society engender greater production of valued economic and social goods than was possible before.[2] In this model, the initial revolutionary drop in output is precipitous ($A' - A$), but the new production curve is much steeper than previously. By point C production is back up to the prerevolutionary level. Still later, the new revolutionary forms have made output higher than would have been attainable without the revolution. Hence, according to the Marxist view, revolution makes up for its preliminary costs, ultimately serving to benefit society.

A third perspective, inspired by Brinton's (1938, chap. 8) work on revolution, is the Thermidorian model. It differs from the previous two by emphasizing neither costs, as the Conservative model does, nor benefits, as the Marxist one does. Rather, the Thermidorian model stresses the return to "normalcy" which occurs after the initial revolutionary losses. At the beginning, there is the inevitable fall in production level (which is overcome by point D). However, prerevolutionary production rates are soon restored, a condition indicated by the absence of a slope change. This view of revolution occupies somewhat of a middle ground, for although development is delayed the society is brought back to health, to draw on Brinton's (1938, pp. 24–26) analogy.

There are, of course, other theoretical possibilities, beyond the Conservative, Marxist, and Thermidorian models diagramed in figure 3. While it seems likely that revolution exacts an initial production cost, one could postulate that the short-run effect is neutral ($\Delta_L = 0$) or even beneficial ($\Delta_L = +$). Further, it may be more realistic to suppose that the initial impact, whatever the direction, would be delayed, rather than immediate; for example, production may fall a year or two after the revolutionary takeover, as parts wear out and managers and technicians leave. In addition, the production curve might not be linear, as assumed. For instance, output over time often follows an exponential path. These alternative model specifications—a nonnegative change in level, a lagged short-run effect, a curvilinear production trend—are entertained later on. Thus, the theoretical structures in figure 3 serve as points of departure.

While the verbal description is, necessarily, sometimes vague or only illustrative, the models themselves provide rather precise rival predictions about revolutionary effects. How can they be tested with actual revolutionary data? Development of a satisfactory statistical model of revolutionary effects requires that the notion of "revolutions as experiments" be

[2] Some analysts may be concerned about spuriousness, worrying that production changes might in fact be caused by productivity changes, rather than by the revolutionary treatment. However, any productivity changes are more appropriately viewed as part of the revolutionary treatment and, as such, could (rightly) account for some of its effect.

thoroughly explored. That is, revolutions ought to be consciously evaluated as social "treatments" with statistically analyzable effects. Following Campbell (1969), who urges this view of reforms, a revolutionary treatment would be no more than quasi-experimental (Cook and Campbell 1976, pp. 223–24). Affirming statistically significant relations can be problematic in the quasi-experimental design of interrupted time series (ITS), which is employed below.

Suppose that within an extended series of annual observations on economic production, a relatively discrete event, a revolution, occurs. I may observe that the level and rate of production appear to drop after the revolution (as the Conservative model would predict). The central statistical question is whether these production differences before and after the revolution are merely due to chance. More formally, are these changes in the level and slope of the production curve statistically significant, thereby ruling out (within acceptable probability limits) the possibility that they are random fluctuations?

This query can be answered through proper estimation of the following linear regression model (Campbell and Cook [1978] propose that an equation of this form generally be used to analyze the effect of intervention in a time series):

$$Y_t = b_0 + b_1X_{1t} + b_2X_{2t} + b_3X_{3t} + e_t\,, \tag{1}$$

where Y_t = yearly observations on economic production; X_{1t} = a counter for years from 1 to N, the number of observations; X_{2t} = a dichotomous variable scored 0 for observations before the revolution and 1 for observations after the revolution; X_{3t} = a counter of years scored 0 for observations before the revolution and 1, 2, 3, . . . for observations after the revolution; b_0, b_1, b_2, b_3 = parameters to be estimated; and e_t = error.

I would like to examine the meaning of the parameters in the model. The parameters b_0 and b_1 estimate, respectively, the level and the slope of the time series before the revolution. The concern here is if the intervention of the revolution affects either of these parameters. Therefore, attention turns to b_2, which estimates the change in level of the series after the revolutionary event, and b_3, which estimates the change in slope after the revolutionary event. The null hypotheses of no change are, respectively, $H_0\colon b_2 = 0$ and $H_0\colon b_3 = 0$. If the coefficient b_2 is not significantly different from zero, one can reject the hypothesis of short-run revolutionary effect. Likewise, if b_3 is not significant, the hypothesis of long-run revolutionary effect can be rejected. These hypotheses are tested in the data analysis which follows.

THE EFFECTS OF THE CUBAN REVOLUTION: SUGAR PRODUCTION

Beyond a doubt, the Cuban Revolution has profoundly changed Cuban society. But there are difficulties in fully quantifying the change, owing to

data problems typical of revolutionary situations and of Third World countries generally. For example, the shortage of observations on health, education, and income distribution before and after the revolution precludes reliable ITS analysis. Omission of these critical categories, in which the revolutionary government is generally acknowledged to have done much, immediately circumscribes my findings. Therefore, I wish to recognize at the outset these important contributions in order to avoid understating the revolution's accomplishments (see Matthews 1975, chaps. 15 and 16). Fortunately, the data limitations are less severe with regard to economic production, where an evaluation of the revolution's impact seems possible. Initially, I consider sugar production and then go on to a more general indicator of Cuban economic performance.

The importance of sugar for Cuba's economy is hard to exaggerate. As far back as 1894, a production of over 1,000,000 tons was reported (Williams and Wright 1975, p. 60). Throughout the 20th century, sugar has dominated Cuban exports and generated a high portion of national wealth. In the period immediately before the revolution (1949–58), sugar made up 84% of the exports and provided 28%–29% of the GNP (Mesa-Lago 1971, p. 278). And, apparently, its impact yet exceeds that demonstrated by these impressive figures. As Henry Wallich explains, the "tendency to pin all hopes on the international sugar markets gives sugar an economic and political dominance even greater than its true weight in the economy" (1950, p. 12). Despite an early effort by the Castro regime to overcome the sugar dependency, it continued to be Cuba's economic backbone (Boorstein 1968, chap. 6; Mesa-Lago 1974, p. 56). Even after failure to achieve the heralded 10,000,000 ton *zafra* in 1970, Castro himself announced: "The fact remains that the nation's basic resource is sugar. . . . And we say again that ours is an economy which depends on one product—sugar cane" (*Granma Weekly Review*, December 20, 1970, pp. 4–5). Further, the leaders of the central planning agency, JUCEPLAN, have indicated that the Cuban economy must concentrate on sugar for some time to come (Matthews 1975, p. 370).

An analysis of sugar output suggests itself not only because of sugar's central economic role, but also because sugar statistics appear more reliable than those on other production and consumption variables, as evidenced by their broad acceptance in the Cuban literature. Nevertheless, Mesa-Lago (1969*a*, p. 59; 1969*b*, pp. 53–54), while accepting the figures up to the mid-1960s, argues that sugar data have been suspicious since 1965, when some inflation may have begun. Such distortions, if they occurred, do not appear to have been serious, for the analysis below reveals no upward bias in the reports from 1965 onward (i.e., actual production was as likely to be below, as above, its predicted value). Beside its strength as a measure, sugar production has the statistical advantage of being the longest available series

($N = 55$), which provides more stable estimates and a greater likelihood of detecting change.

Figure 4 plots annual sugar production (in thousands of metric tons) from 1920 to 1974 (on data sources, see n. 7). The broken line intersecting the series marks the revolutionary takeover in 1959. More precisely, Batista formally governed until December 31, 1958, when he fled the island. Castro's assumption of control dates from January 8, 1959, when he triumphantly entered Havana (Karol 1970, p. 572). Therefore, production reflected in the 1959 observation occurred almost entirely under Castro's revolutionary government, whereas production in 1958, and in the years prior, took place completely under prerevolutionary regimes.

Estimating the ITS model, initially with ordinary least squares (OLS), yields the following:

$$\begin{aligned} Y_t = \underset{(8.93)}{3308.71} + \underset{(2.60)}{43.62}X_{1t} + \underset{(.56)}{408.35}X_{2t} - \underset{(-.34)}{22.73}X_{3t} + e_t, \\ R^2 = .308, \quad d = .838, \end{aligned} \tag{2}$$

where the values below the parameter estimates = the t-ratios; R^2 = the coefficient of multiple determination; d = the Durbin-Watson statistic; Y_t = yearly sugar production (in thousands of metric tons); X_{1t} = years from 1, 2, . . . 55; $X_{2t} = 0$ (before 1959), 1 (1959 and after); $X_{3t} = 0$ (before 1959), 1 (at 1959), 2 (at 1960), . . . 16 (at 1974); and e_t = error.

The coefficients of equation (2) imply that there is no significant effect on sugar production from the revolution, either in the short or long run. However, before accepting this judgment, adjustments should be made for several problems which frequently complicate ITS analysis and which these OLS estimates highlight. The difficulties are as follows: autocorrelation, nonlinearity, and specification of the intervention point. To begin, I address the autocorrelation problem.

Supposing other necessary assumptions are met, OLS will yield desirable estimators when the error terms are uncorrelated, $E(e_ie_j) = 0$, $i \neq j$. For equation (2), $d = .838$, which suggests rather severe autocorrelation difficulties. To adjust for this autocorrelation, I assume a first-order autoregressive model of the error terms, $e_t = \rho e_{t-1} + u_t$, which holds that present error, e_t, is a function of prior error, e_{t-1}, plus some random effect, u_t, that is independent of e_{t-1}. (The appropriateness of this first-order autoregressive assumption and the general problem of autocorrelation in ITS analysis is fully discussed in the footnote.)[8] When error terms are generated by this

[8] I rely on a standard econometric approach in adjusting for autocorrelation. That is, I assume a first-order autoregressive model of the error terms and correct for it. Adoption of a first-order autoregressive model—AR(1) in Box-Jenkins (1970) terminology—is not a serious constraint An AR(1) process dominates annual economic time series, which I have here (Ames and Reiter 1961; Kmenta 1971, p. 278; Pindyck and Rubinfeld 1976, p.

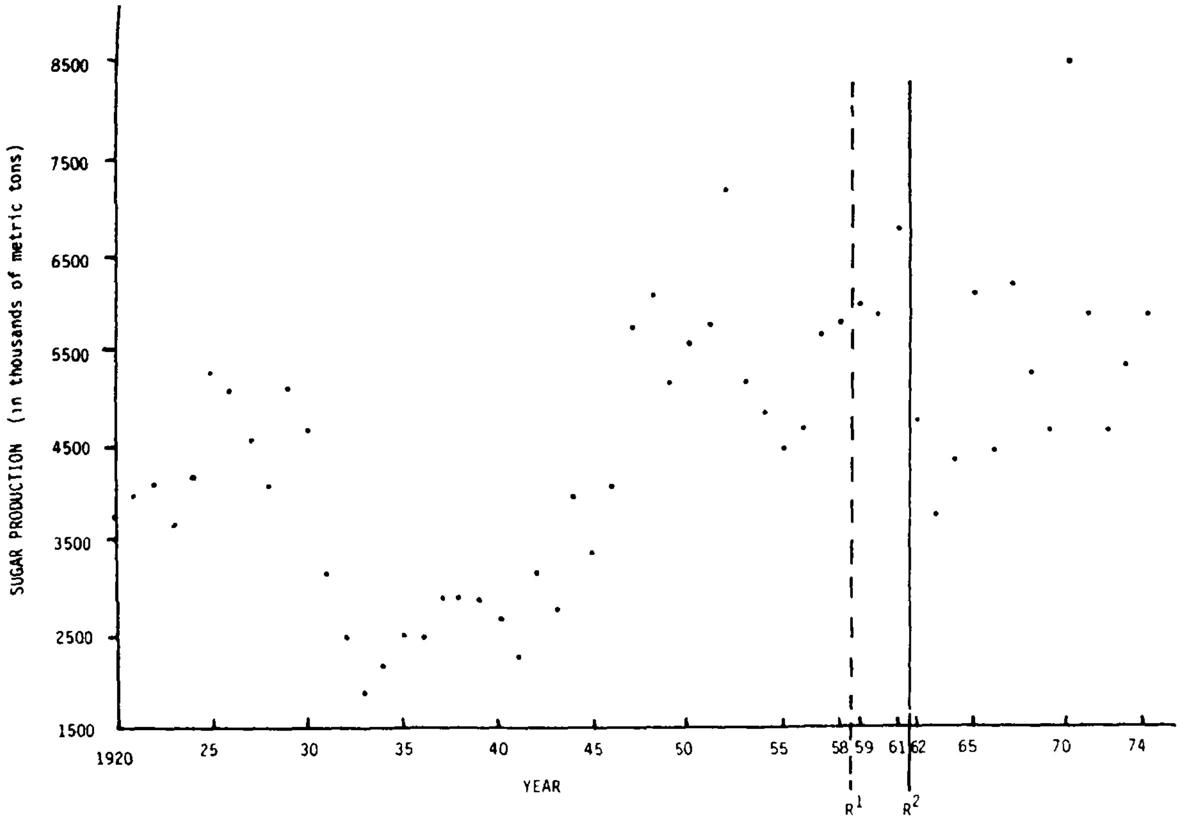


FIG. 4.—Sugar production in Cuba (1920–74)

process, different correction procedures are available. I selected the Cochrane-Orcutt (CORC) technique, which produces desirable estimators and is straightforward and intuitively pleasing. To oversimplify, CORC first estimates the correlation of the error terms, $\rho_{e_t e_{t-1}}$; accordingly, it corrects each observation, subtracting an amount determined by the autocorrelation estimate. These steps are repeated until the estimates converge. By this iterative procedure, estimates equivalent to maximum likelihood estimators are eventually obtained (for a more mathematical explication of the CORC procedure, see Kmenta 1971, pp. 287-89).

The problem of nonlinearity is suggested by a visual inspection of the annual sugar output, which, especially in the earlier years, fluctuated with U.S. prosperity and wartime involvement. Also, it is implied by the modest R^2. Finally, examining the plot of residuals confirms the presence of significant curvilinearity, which can be coped with in various ways. I chose to apply a natural logarithmic transformation to Y_t.

With respect to the specification of the intervention point, the formal transfer of power from Batista to Castro is clearly fixed in time. But, how soon can revolutionary effects be expected to appear? In the Cuban case, as with other revolutions, there was a transition phase. Although exact dates are difficult, 1961 saw the full radicalization of the revolution and the adoption of a "Soviet model" (Ritter 1974, chaps. 3-4). One aspect of this

472). To evaluate this procedure, I examine the corrected level of autocorrelation, using the Durbin-Watson test statistic. (Past research with ARIMA models has employed the Durbin-Watson test as a general evaluation technique [see Nelson 1973, pp. 214-15]) According to the Durbin-Watson value from the Cochrane-Orcutt estimation of eq. (3), autocorrelation ceases to pose a significant threat The autocorrelation problem could be less tractable in other ITS analyses, especially with nonannual, noneconomic observations. Solutions applied in past examples have been inadequate (Campbell and Stanley 1963; Caporaso and Pelowski 1971; Caporaso and Roos 1973; Duvall and Welfling 1973). For certain kinds of time-series analysis, the Box-Jenkins (B-J) forecasting technique offers a flexible procedure for modeling error processes (Box and Jenkins 1970; Hibbs 1974; Nelson 1973). But proposals to apply these B-J methods specifically to ITS analysis often do not fully take into account the difficulties (e g., Campbell and Cook 1978; Glass, Willson, and Gottman 1975). The B-J technique provides a forecasting model, whereas the ITS method examines the causal impact of a real world variable on Y. The causal implications of an ITS design undermine a critical assumption required for B-J procedures, i.e., stationarity of the time series. When a time series, z_t, is stationary, one implication is that the expected values of any two observations are the same, $E(z_t) = E(z_{t+m})$. Therefore, the mean of a series should be a constant, independent of time (see Nelson [1973, pp. 19-23] for a useful discussion of stationarity) The dilemma for an ITS design becomes apparent immediately. When an intervention has an effect (e.g., an intercept change), the mean of the series should shift from a pretreatment to a posttreatment level, rendering the series nonstationary. (Further, this sort of nonstationarity is not homogeneous and cannot be removed by differencing.) These particular difficulties are just beginning to be considered in the statistical literature, and significant practical problems of modeling and estimation remain (see Box and Tiao [1975] and Hibbs [1977*a*, 1977*b*], who has advanced the work here significantly). Finally, as Campbell and Cook (1978) themselves admit, economists have shied from a B-J approach because of its treatment of the stationarity question.

economic restructuring was a de-emphasis on sugar, as the revolutionary leadership sought to transform Cuba from an agricultural to an industrial state. The precipitous drop in sugar output from 1961 to 1962, shown in figure 4, suggests the consequences of this fundamental reorientation. Further, it argues that the short-run revolutionary impact is more fully observed in 1962, rather than in 1959 (see the solid line, fig. 4).

In view of the above considerations, I use the CORC procedure to re-estimate the ITS equation, after having logged sugar production (to the base *e*) and respecified the short-run effect hypothesis (hence, for X_{2t} and X_{3t}, 1962 = 1):

$$\log_e Y_t = 7.885 + .019X_{1t} - .328X_{2t} + .009X_{3t} + e_t, \quad (3)$$
$$(29.24) \quad (1.92) \quad (-1.54) \quad (.24)$$
$$R^2 = .678, \quad d = 2.197.$$

These estimates, arrived at after two iterations, reflect adjustments for the high autocorrelation that affected this series, $\hat{\rho}_{\hat{e}_t\hat{e}_{t-1}} = .737$. The Durbin-Watson statistic indicates that autocorrelation is no longer a problem. Further, nonlinearity has ceased to be troublesome. Besides the improvement in "fit," an inspection of the residual plot reveals no significant curvilinearity (i.e., the residuals are more or less randomly scattered across the time span).[4] These tests encourage acceptance of the estimates in equation (3) (further assurance comes from ruling out a possible multicollinearity threat).[5]

[4] In general, the closer the Durbin-Watson statistic, *d*, is to 2, the better. But the precise value for *d* at which the null hypothesis of no autocorrelation is accepted varies with the level of significance, whether the test is one- or two-tailed, the sample size, and the number of independent variables. Here I have three independent variables, a 5% level of significance with a two-tailed test, and $N = 54$ (one of the original cases is necessarily lost with the CORC adjustments). Under these conditions, if the value of *d* lies between 1.60 (d_u) and 2.40 ($4 - d_u$), as it does in this case, then the null hypothesis of no autocorrelation can be accepted (see Kelejian and Oates 1974, pp. 200–3, 278). Interestingly, with an ITS design, the Durbin-Watson test serves as a check on nonlinearity as well as autocorrelation. One test of linearity is whether residuals, when arrayed according to *X*, are randomly scattered around the regression line. If *X* is a counter of time, as with ITS, then the Durbin-Watson test, by indicating that errors are uncorrelated, also suggests the absence of nonlinearity (see discussion in Kmenta 1971, pp. 470–71). Of course, the implication is that an autocorrelation correction procedure, such as CORC, will tend to remove nonlinearity as well.

[5] While a multicollinearity problem is suggested, a rather demanding test shows that there is none. The correlation between X_{2t} and X_{3t} for the sugar production series is high, $r = .85$ (a coefficient of this magnitude comes from the very large number of prerevolutionary scores here for X_{2t} and X_{3t}, which the ITS model sets at zero). This implies that the estimates of eq. (3) could be wildly off, because the effects of X_{2t} and X_{3t} are not being reliably separated. When multicollinearity is operating in this way, solution is problematic. One approach is to analyze the effects of X_{2t} and X_{3t}, in turn, as follows:

$$\log_e Y_t = 7.854 + .021X_{1t} - .310X_{2t} + e_t$$
$$(32.15) \quad (2.53) \quad (-1.59)$$

Although the coefficient for the short-run effect, b_2, is in the expected direction, it is not statistically significant (at the .05 level, for a one-tailed test, $|t|$ must exceed 1.65). Thus, responsibility for a sugar production drop from 1961 to 1962 cannot be confidently attributed to the revolutionary government. Possibly, the decline would have occurred under the pre-revolutionary government as well. (Also, as expected, no short-run effects for the revolution emerge when the other possible intervention points are analyzed, i.e., when for X_{2t} and X_{3t}, 1 = 1959, 1960, 1961, or 1963, b_2 does not approach statistical significance.) With regard to the long-run impact of the revolution, although b_3 is positive, the t-ratio indicates that it is far from statistical significance. (Likewise, when other intervention points—1959, 1960, 1961, 1963—are examined, the revolution does not come close to achieving a statistically significant long-term effect.)

In his recent book on Cuba, economist Mesa-Lago singles out sugar, asserting that it "has had a discouragingly bad performance under the Revolution" (1974, p. 56). While this evaluation is more nearly correct for the short run (1962), it certainly appears false for the long run. Indeed, if I were to generalize, in the fashion of Mesa-Lago, I would have to conclude that the revolution has had no effect on sugar production. What is more surprising, however, is the degree to which sugar production levels resist change, even in the face of serious government initiatives. This finding has profound implications for the formulation of Cuban development policy. It suggests that the best efforts of any Cuban government can heighten sugar output only marginally. Sugar production capacity may have reached a point where it is highly constrained by factors that the government cannot greatly alter, such as the island's agricultural and human resource base, not to mention international patterns of trade and dependency.

These findings fail to support any of the models of effect postulated earlier, although equation (3) tends to the Thermidorian prediction. However, the evidence on sugar may mislead for, while very important, sugar represents only one component of Cuban economic production. To speak with some assurance about the revolution's global impact on economic growth, it is necessary to analyze a more aggregate, GNP-type indicator.

---

and

$$\log_e Y_t = 7.947 + .016X_{1t} - .012X_{3t} + e_t .$$
$$(31.30) \quad (1.68) \qquad (-.38)$$

The revolution still manages to exhibit no statistically significant effects, even allowing each intervention variable to explain all of Y_t it can, free of the control of the other. Furthermore, their t-ratios and parameter estimates are quite close to those in the full ITS equation. Therefore, I conclude that multicollinearity does not pose a serious threat to the interpretation of the eq. (3) results.

THE EFFECTS OF THE CUBAN REVOLUTION: ENERGY CONSUMPTION AND ECONOMIC GROWTH

While there is broad recognition of the revolution's beneficial social consequences, for example, improved social services, medical care, and educational opportunities, sharp disagreement exists about its overall effect on economic production. Among those who contend that the revolution has had a favorable economic influence, a "stagnation thesis" is popular (Boorstein 1968, chap. 1; Matthews 1975, pp. 35-39, 292-318; O'Connor 1970, pp. 17-18; Roca and Hernandez 1972, p. 68; Seers 1964, pp. 3-19). According to this perspective, prerevolutionary Cuba had shown almost no signs of economic growth as far back as the 1920s. The revolution eliminated the patterns that produced this stagnation and initiated a period of lasting, significant economic advance. But there are others who claim, for a variety of reasons, that the revolution has not had an entirely salubrious effect on Cuba's economic development (Dumont 1970; Gonzalez 1974, pp. 193-218; Karol 1970, pp. 407-84; Nelson 1972). One of Castro's strongest critics, Lowry Nelson, states his opinion quite flatly: "Cuba constitutes a unique example of a country launched on a classical linear path of capitalist development, a course that is suddenly truncated by revolution. . . . The consequence to Cuba has been an almost unmitigated economic disaster" (1972, pp. 36, 196). My analysis does not fully support any of these opposing generalizations.

To respond to the question of the revolutionary impact on economic growth, the focus should be on indicators of central tendencies in the Cuban economy. Of course, GNP figures would be ideal. Unfortunately, in addition to gaps in the series, the change in computation of national accounts from capitalist to socialist techniques rules out reliable comparison with prerevolutionary figures (Mesa-Lago 1969*b*, pp. 48-51; 1971, pp. 323-24). Faced with such a dilemma, one must seek indicators which are serviceable proxies for GNP.

David McClelland, in his imaginative work *The Achieving Society*, decided that electric energy production reliably indicated a nation's economic production (1961, pp. 82-87). McClelland selected electricity output as an indicator because he lacked adequate national income figures, as is also the case here. He reasoned that most of the productive machinery of contemporary societies would not run without electrical power (pp. 85-87). Therefore, electric output, as measured in kilowatt hours, would be an indicator of a nation's material production. This rather novel measurement appears to have substantial empirical validity, as evidenced by a correlation of .865 between per capita electricity production and the Watkins-Hagen estimates of per capita national income for a sample of 40 countries (McClelland 1961, p. 87). Hence, using electrical energy data to represent GNP, it may be possible to assess indirectly the revolution's contribution to

Cuban economic growth. A series on electrical energy production can be assembled. However, the observations represent only public production, not total production (Mesa-Lago 1969*b*, p. 59). Since nationalization by the revolutionary government expanded the state sector, these figures would at least partially reflect the change to socialism, rather than real economic growth. But, there are also data on energy consumption, which McClelland equates with energy production (1961, p. 86). These figures on energy consumption include all consumption, public or private, and are thus not subject to the shortcomings of the energy production figures.

In the absence of adequate national accounts data, energy consumption is not a totally unreasonable proxy. Besides the empirical overlap with national income, indicated by McClelland, it has the advantage of being an unobtrusive measure. Energy consumption figures seem much less likely to be "fudged" than GNP figures, which easily lend themselves to propaganda purposes. Critic Mesa-Lago (1969*b*, p. 58) appears to accept electric energy statistics as accurate, thereby giving support to this inference. Further, an ITS analysis will resist various common sources of error.[6]

In addition to the foregoing arguments, the Cuban data themselves suggest that the energy consumption variable is tapping GNP. First, energy consumption is highly correlated with production in key Cuban industries for which data are available: with cement production, $r = .84$; cigarette production, $r = .72$; crude petroleum production, $r = .77$; motor fuel production, $r = .70$; kerosene production, $r = .94$; phosphate fertilizer consumption, $r = .80$; nitrogenous fertilizer consumption, $r = .86$; potash fertilizer consumption, $r = .79$.

More general evidence comes from a principal components analysis of all the production and consumption variables (23) on which I could assemble a more or less adequate series.[7] While the first factor accounts for 53% of

[6] Under a condition of constant error, e.g , figures from the Castro and Batista regimes are equally inflated, ITS estimation of significant pretreatment to posttreatment change remains unaffected (Costner 1971, pp. 299–303). Mesa-Lago himself argues that regimes as opposing as those of Castro and Batista tend to generate the same biases in statistical information (1969*b*, p. 75).

[7] An interrupted time series must have "enough" pretreatment and posttreatment observations to permit meaningful application of the necessary statistical tests. The average series length is 37 observations for these Cuban data, with the last observation always from 1973 or 1974. With regard to missing data, the aim was to include just those series which contained no gaps (minor interpolation of one or two values was necessary for beer production, cigarette production, nitrogenous fertilizer consumption, motor fuel production, kerosene production, and phosphate fertilizer consumption). In addition to these, series were also gathered on the following production variables: beans, cement, chromium ore, cocoa, coffee, copper ore, corn, crude petroleum, electric energy, peanuts, rice, salt, sugar, and tobacco; and on these consumption variables: energy, newsprint, and potash fertilizer. These total to 23 variables. The agricultural production data were gathered from volumes of the following statistical yearbooks: *International Yearbook of Agricultural Statistics* (Rome: Institut International d'Agriculture, 1922–47); *Yearbook of Food and*

the variance in the data, the remaining factors individually contribute little (e.g., the second factor adds only 10% more to the variance explained). Perhaps this first factor, which represents the best linear combination of all 23 variables, is a partial measure of national production. Interestingly, the variable having the highest factor loading is energy consumption, with .97.[8] This extremely high loading indicates that, at least empirically, the energy consumption variable is equivalent to this production factor. Thus, I have some confidence that energy consumption is a workable substitute for GNP.

In figure 5, 26 annual observations on energy consumption from 1949 to 1974 are plotted (the values are natural logarithms of actual consumption, measured in millions of metric tons; data sources are described in n. 7) The revolutionary intervention is indicated by the broken vertical line between 1958 and 1959.

Estimating the ITS model with OLS yields

$$\log_e Y_t = .836 + .076X_{1t} + .089X_{2t} - .032X_{3t} + e_t,$$
$$(35.35) \quad (20.06) \quad (3.26) \quad (-7.43) \qquad (4)$$
$$R^2 = .995, \quad d = 1.702.$$

According to these OLS estimates, there are significant changes at the .05 level (two-tail, $|t| > 2.06$) in the level and slope of the energy consumption curve after the revolution, as indicated by b_2 and b_3. Further, the Durbin-Watson statistic indicates the absence of autocorrelation, allowing the acceptance of these significant results as unbiased. A diagram nicely illustrates the meaning of these changes and the workings of the ITS design. I would like to consider the slope change first. In figure 5, the predicted consumption values from the above multiple-regression equation are linked together, first for the preintervention years and then for the postintervention years. The result is two straight lines, which actually provide simple linear regression fits to the points within the particular time period. Visually, the slopes of the two lines appear to differ. In fact, the slope of the earlier line is .076 (b_1), while the slope of the later line is only .044. This drop in slope of .032 is indicated by b_3 and is statistically significant.

Agricultural Statistics, pt. 1 (Rome: Food and Agricultural Organization of the United Nations, 1949–57); *Production Yearbook* (Rome Food and Agricultural Organization of the United Nations, 1958–74). The remaining data were gathered from volumes of the *United Nations Statistical Yearbook* (New York. Statistical Office of the United Nations, 1948–75).

[8] This energy consumption loading is on the first factor of a principal component analysis (unrotated) of the above 23 variables (logged to the base e) This loading remains almost exactly the same, regardless of whether the variables are unlogged, the factors are rotated (varimax), or the energy production variable is excluded Further, the percentage of variance explained by the first factor is stable under these different conditions.

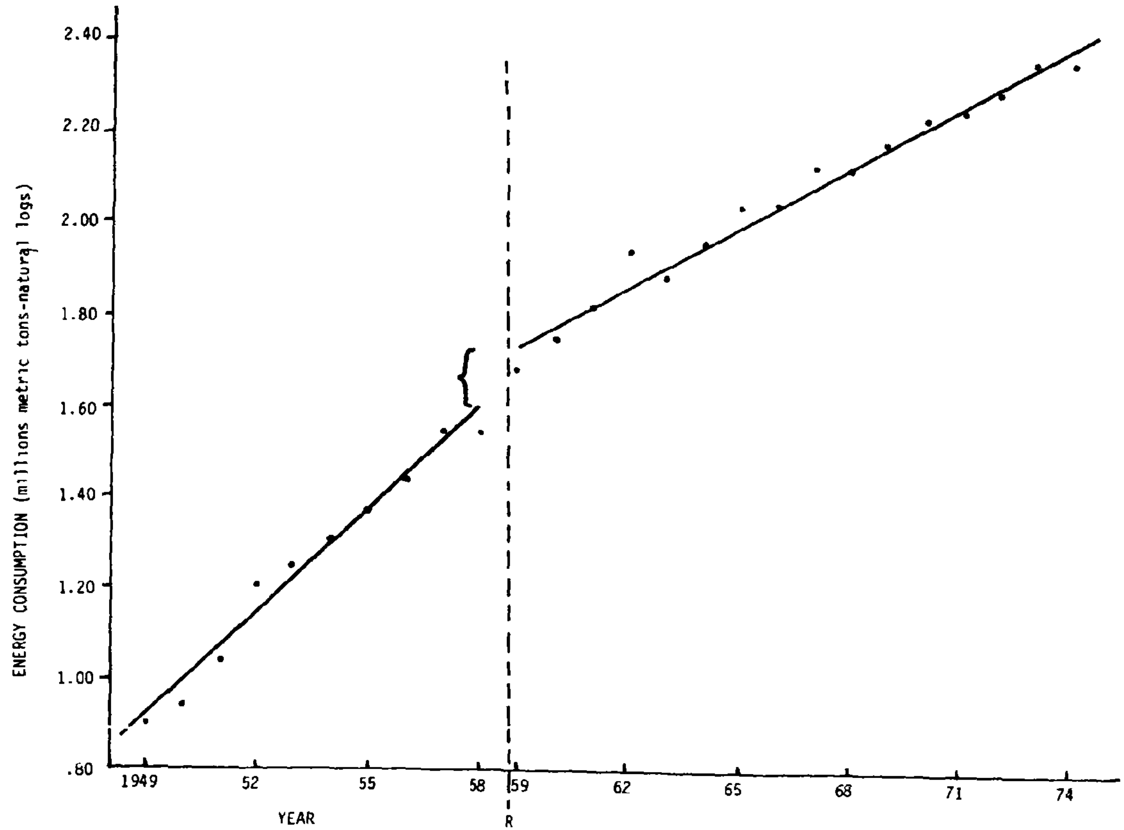


FIG. 5.—*Interrupted time-series analysis of Cuban energy consumption (1949–74)*

Turning to the level, the change here is suggested by the breaking of the two lines between the 1958 and the 1959 values. For the prerevolutionary years, the prediction equation reduces to $\log_e \hat{Y} = b_0 + b_1X_{1t}$ (because X_{2t} and X_{3t} take on zero values). Supposing the revolution had not occurred, the prediction for 1959 based on this prerevolutionary model would be 1.6763. However, the revolution did occur, and the corrected prediction for 1959 becomes 1.7337, showing an increase of .0574. This increase is caused by the significant postintervention intercept rise (b_2) and would be even greater if the depressing effect of the negative slope change (b_3) were not operating ($b_2 - b_3 = .0889 - .0315 = .0574$).

As noted, the observations on energy consumption are expressed in "natural" logarithms (logs to the base e), while the independent variables, which are various measures of time, remain unlogged. This semilog transformation is especially useful when the concern is with assessing rates of economic growth. It tends to "straighten out" the observations, which in their raw form usually follow a curve that steepens over time. This linearization, of course, makes an OLS analysis more appropriate. But, in addition, the transformation allows the evaluation of percentage change, rather than absolute change. For example, if OLS were used to estimate a simple exponential model predicting energy consumption from time (in years), $\log_e Y_t = b_0 + b_1X_{1t} + e_t$, the interpretation would be that a one-year increment yields a ($b_1 \times 100$) percentage increase in energy consumption.[9]

With this understanding, the ITS estimates for energy consumption in equation (4) can be more fully considered. The slope, b_1, estimates that energy consumption grew at a rate of 7.6% a year during the prerevolutionary period. As b_3 indicates, after the revolution this rate was dampened by 3.2%, a significant decline. Therefore, under Castro the estimated annual energy consumption rate stands at 4.4%. How do these estimates compare with those of GNP growth? Ritter, in response to the serious data limitations, remarks: "Research on the Cuban economy is thus reduced to a type of detective work" (1974, p. vii). I proceed in this spirit.

The only reliable national accounts statistics for the prerevolutionary period, according to Mesa-Lago (1969*b*, pp. 48–49), came from the National Bank, which reported an annual GNP growth rate in current prices for 1950–58 of 4.6%. However, the preferred comparison would actually seem to be with GNP growth in constant prices, since energy consumption is measured in units of constant value (i.e., metric tons). Although the National Bank did not report GNP in constant prices, consumer price indices from Havana implied an inflation of 1.1% a year (Mesa-Lago 1971, p. 278). This figure suggests that the annual growth rate, in constant prices, was around 3.5% for the period. Accepting this estimate, one might infer a

[9] A proof for this interpretation, which becomes problematic as b_1 exceeds .25, is found in Tufte (1974, pp. 124–26).

constant average positive error of 4.1% a year (7.6% − 3.5%) when using energy consumption growth to estimate real GNP growth. (This supposition is not entirely arbitrary. Even if GNP and energy consumption were perfectly correlated, the energy consumption–GNP conversion might well be inefficient.) Therefore, GNP growth, at constant prices, under the revolutionary government would be more correctly estimated at .3% annually.

However, this estimate in some sense exaggerates economic growth, for it does not control for population increase. Therefore, I reestimate (OLS) the ITS equation, but with per capita energy consumption (logged to the base e) as the dependent variable:

$$\log_e Y_t = \underset{(253.95)}{6.115} + \underset{(13.33)}{.052X_{1t}} + \underset{(3.17)}{.088X_{2t}} - \underset{(-6.55)}{.028X_{3t}} + e_t, \quad (5)$$

$$R^2 = .987, \quad d = 1.720.$$

While the pattern remains the same as in equation (4), the coefficients are reduced, as would be expected. Beginning with the National Bank statistics noted above, the growth of GNP per capita for 1950–58, in constant prices, can be estimated at 1% annually (the period's annual population growth rate of 2.5% subtracted from the real GNP rate [see Mesa-Lago 1971, p. 278]). If this estimate is correct, it suggests an average constant positive error of 4.2% a year (5.2% − 1.0%) in using per capita energy consumption growth to predict real per capita GNP growth. Applying this adjustment for the Castro period yields a real annual per capita GNP growth rate estimate of −1.8% (revolutionary growth rate = [5.2% − 2.8%] − 4.2% = −1.8%).

Comparison with more standard, although less comprehensive, analyses suggests that this estimate of economic growth under the revolution is not completely fanciful. United Nations figures for 1962–66 indicate that real GNP per capita fell .5% per annum (Mesa-Lago 1969*b*, p. 5). For the period 1960–70, the World Bank reports a real GNP per capita growth rate of −.6% a year (World Bank Group 1972, p. 3). Another evaluation, for 1960–69, shows a real GNP per capita decrease of 3.2% yearly (International Monetary Fund 1972, p. 51). Ritter, however, argues convincingly that this last estimate is probably an exaggeration (1974, p. 209). On the basis of these investigations, my estimate of an annual 1.8% decline in real GNP per capita for the entire period, 1959–74, would not appear wide of the mark. Its credibility is enhanced further upon consideration of the meaning of the significant intercept shift (b_2).

For the period of revolutionary transformation, there is rather broad agreement that economic performance, especially in 1959, was better than in the prerevolutionary period (Mesa-Lago 1969*b*, p. 51; Ritter 1974, p. 106). However, because there is an unusually serious lack of economic data from

this time span, it has been difficult to specify precisely the degree of improvement. Available guesses focus on the annual rate of GNP growth in current prices. Estimates range from 12% annually, by Boti, for 1959–60 to 6% annually, by Effimov, for 1959–61 (see Mesa-Lago 1969*b*, p. 50). Noyola (1961) provides estimates of 10.5% a year for both 1959 and 1960. However, Pazos (1962) argues that these figures are too high, proposing a 1959 figure of from 5% to 6%.

Recalling the explanation of the level change for the energy consumption curve (see the discussion of fig. 5), the estimated energy consumption growth rate for 1959 would be 13.3% (from eq. [4], $b_1 + [b_2 - b_3] = .076 + [.089 - .032] = .076 + .057 = .133$). If this prediction were corrected downward by 4.1% as done earlier, then the estimated 1959 GNP growth rate, in constant prices, = 9.2% (adding the 1.1% inflation figure yields a 1959 GNP growth estimate, in current prices, of 10.3%, which is almost exactly Noyola's estimate). But, again, a more accurate growth picture probably comes from the per capita energy consumption equation prediction of the real per capita GNP growth. In this case, the 1959 estimate for the real per capita GNP growth rate = 7.0% (from eq. [5], $b_1 + b_2 - b_3$ − constant error $= .052 + .088 - .028 - .042 = .070$). Hence, this ITS formulation of the energy consumption trend is correctly detecting the revolutionary high point in economic growth postulated by careful observers of the Cuban economy. Moreover, the fullest short-run revolutionary effect is captured when the first postintervention observation is 1959, rather than a later year; therefore, the initial model specification for X_2 (and X_3) in equations (4) and (5) is sustained.[10]

THEORETICAL IMPLICATIONS

These findings support none of the rival models—Conservative, Marxist, and Thermidorian—developed originally. The rejection of these three models, all of which share the assumption of short-run revolutionary costs, implies that the initial economic effects of revolutions, at least in the Cuban case, may not be as harmful as generally supposed Rather, with regard to economic growth, short-run benefits and long-run costs appear to follow revolution. Metaphorically, the pattern might be labeled "euphoric," with elation replaced by gloom. For Cuba in 1959, the per capita real GNP

[10] In altering the point of intervention (for X_2 and X_3, year 1 = 1959, 1960, 1961, 1962, or 1963), one observes that b_2 declines steadily and ceases to be statistically significant for year 1 = 1962 or 1963. In contrast, the estimates for b_1 and b_3 remain very stable in the face of these manipulations. This stability heightens confidence in the reliability of the estimates of the per capita energy consumption rates for the pre- and postperiods, for it indicates that they are not merely a product of the particular starting and stopping points in eq. (5), i.e., 1949–58, 1959–74. (This same stability of b_1 and b_3 is observed when non–per capita energy consumption is so analyzed; likewise, the estimate for b_2 declines in a similar way, ceasing to be statistically significant if year 1 = 1962 or 1963.)

growth rate was 7.0%, by my estimate. This preliminary production gain came largely from the increased demand made possible by income redistribution and from the greater use of already existing resources. However, these economic advances were shortly negated by trade difficulties, the scarcity of managers and materials, and administrative error. In fact, looking across the entire Castro period, I estimate the annual rate of real per capita GNP growth at −1.8%, compared with 1.0% for the prerevolutionary era.

Nevertheless, even granting that the economic growth rate has fallen during the revolutionary years, the inference that Castro's policies are responsible for the decline is hasty. While the Cuban government obviously directs the Cuban economy, there are external influences over which it has little control. Of overwhelming importance here is U.S. policy. Specifically, what has been the impact of the U.S. embargo? In the case at hand, the imposition of the embargo is almost coincident with the revolution and so may share responsibility for the drop in economic output. Ideally, revolutionary effects should be tested in the absence of an embargo.

This ideal circumstance prevailed during the first year of the revolution. Throughout 1959, international trade relations followed the prerevolutionary pattern (serious U.S. economic reprisals began in 1960 with reduction of the sugar quota in July and the imposition of a partial embargo in October). Thus, a partial test of revolutionary effects, disentangled from the embargo, is found in the short-run effect coefficient, b_2, of eq. (5). As discussed, in this first year growth was significantly greater than for 1958. Also, sugar production held steady, with no indication of a drop from the previous year ($b_2 = .031$, $t = .146$). However, the embargo soon led to serious shortages of raw materials and replacement parts for industry, and output slowed. Hence, the temptation is to attribute the long-run production decrease to persisting effects of the embargo. But, at about the same time the embargo was taking hold, important internal changes in Cuba were occurring, for example, the transition to a socialist state was being completed, and the Soviet Union was assuming a major role. Therefore, by the 1960–61 period, major variables potentially influencing growth are once again confounded. Nevertheless, the brief 1959 experiment hints at the revolutionary possibilities for positive economic growth. Without a doubt, the subsequent decline was at least partially caused by the far-reaching economic reorientation that the U.S. sanctions required and cannot fairly be ascribed solely to the Castro government.

To what extent can one expect the Cuban findings to hold for other revolutions? Is the impact on Cuba different from the impact on Russia? Or France? Perhaps, as the Cuban case suggests, the consequences of revolution vary according to the nation's size and international dependency. It may be that its effects are associated with the initial economic development level of the country or with the social origins of the revolutionary

leadership. To the degree that such conditions pertain, unqualified inferences from the Cuban experience to the universe of revolutions are unwarranted. Clearly, in building a theory of effects, more revolutions must be studied, and the implication is that statistical interactions must be seriously considered. Further, investigation of other revolutions promises longer time series, and the opportunity to include intermediate- as well as short- and long-run effects in an ITS analysis.

Revolutionary effects may be conditioned by the dependent variables themselves, as well as by certain independent variables. In my investigation, I focused on the impact of revolution on economic variables. But it is possible, even likely, that revolution tends to influence social and political variables differently. Thus, while additional research might provide further support for a "euphoric" model with regard to economic growth, that model might not hold in other areas of society. One of the three models posed at the outset might actually apply. Or yet another formation could be relevant. For example, with regard to social variables, the typical pattern may be a sort of "mobilization" model (e.g., educational attainment levels are at first unaffected by the revolution [$b_2 = 0$] but then commence a steady rise [$b_3 > 0$]).

Entertaining the idea of different models for different sets of dependent variables throws into sharp relief the complexity of the perennial question of social philosophers, "Is revolution worthwhile?" Even if this question could somehow be answered for each sector—the economic, the social, and the political—it would persist at a higher level. For instance, what would be the objective criteria for determining that the social benefits of revolution outweighed the political costs? Premier Castro, in a 1971 speech in Chile, declared that "the accomplishments of a revolution are not to be measured only in stones; they are not to be measured only in factories. They *are* to be measured in these, but essentially they ought to be measured in moral and human terms" (as quoted in Matthews 1975, p. 439). I share Castro's concern that all the outcomes of revolution, the moral along with the material, be measured. But evaluating the worth of a revolution ultimately reaches beyond measurement of the relevant variables. Even after all the various revolutionary outcomes are accurately recorded, it remains to be asked, "Which do I prefer?" and "How much do I prefer them?" These questions, of course, force us to look beyond the methodology of this paper.

REFERENCES

Ames, E., and S. Reiter. 1961. "Distributions of Correlation Coefficients in Economic Time Series." *Journal of the American Statistical Association* 56 (September): 637–56.

Boorstein, Edward. 1968. *The Economic Transformation of Cuba*. New York: Monthly Review Press.

Box, G. E. P., and G. M. Jenkins. 1970. *Time-Series Analysis: Forecasting and Control*. San Francisco: Holden Day.

Box, G. E. P., and G. C. Tiao. 1975. "Intervention Analysis with Applications to Economic and Environmental Problems." *Journal of the American Statistical Association* 70 (March): 70–79.

Brinton, Crane. 1938. *The Anatomy of Revolution*. New York: Norton.

Campbell, Donald T. 1969. "Reforms as Experiments." *American Psychologist* 24 (April): 409–29.

Campbell, Donald T., and Thomas D. Cook. 1978. *The Design and Analysis of Quasi-Experiments for Field Settings*. Chicago: Rand McNally.

Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Caporaso, James A., and A. L. Pelowski. 1971. "Economic and Political Integration in Europe: A Time-Series Quasi-Experimental Analysis." *American Political Science Review* 65 (June): 418–33.

Caporaso, James A., and Leslie J. Roos, Jr., eds. 1973 *Quasi-Experimental Approaches: Testing Theory and Evaluating Policy*. Evanston, Ill.: Northwestern University Press.

Cook, Thomas D., and Donald T. Campbell. 1976. "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings." Pp. 223–326 in *Handbook of Industrial and Organizational Psychology*, edited by Marvin D. Dunnette. Chicago: Rand McNally.

Costner, Herbert L. 1971. "Theory, Deduction, and Rules of Correspondence " Pp. 299–319 in *Causal Models in the Social Sciences*, edited by Hubert M. Blalock, Jr. Chicago: Aldine.

Davies, James C. 1962. "Toward a Theory of Revolution." *American Sociological Review* 27 (February): 5–19.

Dumont, René. 1970. *Cuba Est-il Socialiste?* Paris: Editions du Seuil.

Duvall, R., and M. Welfling. 1973. "Determinants of Political Institutionalization in Black Africa: A Quasi-Experimental Analysis." *Comparative Political Studies* 5 (January): 387–417.

Eckstein, Harry. 1964. "Introduction: Toward a Theoretical Study of Internal War." Pp. 1–31 in *Internal War: Problems and Approaches*, edited by Harry Eckstein. New York: Free Press.

Glass, G. V., V. L. Willson, and J. M. Gottman. 1975. *Design and Analysis of Time-Series Experiments*. Boulder: Colorado Associated University Press.

Gonzalez, Edward. 1974. *Cuba under Castro: The Limits of Charisma*. Boston: Houghton Mifflin.

Greene, Thomas H. 1974 *Comparative Revolutionary Movements*. Englewood Cliffs, N.J.: Prentice-Hall.

Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, N.J.: Princeton University Press.

Hibbs, Douglas A. 1974. "Problems of Statistical Estimation and Causal Inference in Time-Series Regression Models." Pp. 252–308 in *Sociological Methodology, 1973–1974*, edited by H. L. Costner. San Francisco: Jossey-Bass.

———. 1977*a*. "On Analyzing the Effects of Policy Interventions: Box-Jenkins and Box-Tiao vs. Structural Equation Models." Pp. 137–79 in *Sociological Methodology, 1977*, edited by D. Heise. San Francisco: Jossey-Bass.

———. 1977*b*. "Political Parties and Macroeconomic Policy." *American Political Science Review* 71 (December): 1467–87.

Huntington, Samuel P. 1968. *Political Order in Changing Societies*. New Haven, Conn.: Yale University Press.

International Monetary Fund. 1972. "Gross National Product per Capita (1969) and Average Annual Growth Rate (1960–1969)." *Finance and Development* 9 (March): 51.

Johnson, Chalmers. 1966. *Revolutionary Change*. Boston: Little, Brown.

Karol, K. S. 1970. *Guerillas in Power: The Course of the Cuban Revolution*. New York: Hill & Wang.

Kelejian, Harry H., and Wallace E. Oates. 1974. *Introduction to Econometrics: Principles and Applications*. New York: Harper & Row.

Kmenta, Jan. 1971. *Elements of Econometrics*. New York: Macmillan.

McClelland, David C. 1961. *The Achieving Society*. New York: Free Press.

Matthews, Herbert L. 1975 *Revolution in Cuba*. New York· Scribner's.

Mesa-Lago, Carmelo 1969*a* and 1969*b* "Availability and Reliability of Statistics in Socialist Cuba." *Latin American Research Review* 4 (Spring): 53-91; and 4 (Summer)· 47-82.

———. 1971. "Economic Policies and Growth." Pp 277-338 in *Revolutionary Change in Cuba*. Pittsburgh: University of Pittsburgh Press

———. 1974. *Cuba in the 1970's: Pragmatism and Institutionalization* Albuquerque: University of New Mexico Press.

Moore, Barrington, Jr. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World* Boston Beacon

Nelson, Charles 1973. *Applied Time Series Analysis*. San Francisco: Holden-Day

Nelson, Lowry 1972. *Cuba: The Measure of a Revolution* Minneapolis University of Minnesota Press

Noyola, J. 1961. "La revolución Cubana y sus efectos en el desarrollo económico." *El trimestre económico* 28 (July-September): 416

O'Connor, James. 1970 *The Origins of Socialism in Cuba* Ithaca, N Y . Cornell University Press.

Pazos, F. 1962. "Comentarios a dos articulos sobre la revolución Cubana." *El trimestre económico* 29 (January-March): 10-13.

Pindyck, Robert S., and Daniel L. Rubinfeld. 1976. *Econometric Models and Economic Forecasts*. New York: McGraw-Hill

Ritter, Archibald R. M 1974. *The Economic Development of Revolutionary Cuba: Strategy and Performance*. New York. Praeger

Roca, Sergio, and Roberto E. Hernández 1972 "Structural Economic Problems." Pp. 67-93 in *Cuba, Castro, and Revolution*, edited by Jaime Suchlicki Coral Gables, Fla.· University of Miami Press

Seers, Dudley. 1964 "The Economic and Social Background" Pp 3-61 in *Cuba: The Economic and Social Revolution*. Chapel Hill: University of North Carolina Press.

Smelser, Neil J. 1963. *Theory of Collective Behavior*. New York. Free Press.

Tufte, Edward R. 1974 *Data Analysis for Politics and Policy* Englewood Cliffs, N J : Prentice-Hall.

Wallich, Henry C. 1950. *Monetary Problems of an Export Economy: The Cuban Experience, 1914-1947*. Cambridge, Mass.: Harvard University Press.

Williams, Edward J., and Freeman J. Wright. 1975. *Latin American Politics: A Developmental Approach*. Palo Alto, Calif : Mayfield.

Wolf, Eric R. 1969. *Peasant Wars of the Twentieth Century* New York: Harper & Row.

World Bank Group. 1972. *World Bank Atlas, 1972: Population, Per Capita Production Growth Rates*. Washington, D.C.: World Bank.

Suicide, Motor Vehicle Fatalities, and the Mass Media: Evidence toward a Theory of Suggestion[1]

David P. Phillips
University of California, San Diego

Tarde and other classical sociologists paid a great deal of attention to the concepts of imitation and suggestion, but these concepts have been virtually ignored in modern sociology. This paper presents new findings indicating that imitation and suggestion have a powerful impact on social behavior: Three days after a publicized suicide, automobile fatalities increase by 31%. The more the suicide is publicized, the more the automobile fatalies increase. The age of the drivers is significantly correlated with the age of the person described in the suicide story. Single-car accidents increase more than other types just after the publicized suicide. After presentation of these and related findings, the paper discusses some ways in which the concepts of suggestion and imitation can be incorporated into sociological theory.

Contemporary sociologists have largely neglected the social effects of suggestion, modeling, and imitation.[2] This neglect is somewhat puzzling in view

[1] The research reported here was supported by grant R-64 from the University of California Regents. I thank Mark Murphy, city editor of the *Los Angeles Times*, for granting me access to the *Los Angeles Times* reference library and for helping me to gain access to the reference library of the *San Francisco Chronicle;* Grace Hazzard and Reuben Harding, Statistical Information Staff, National Highway Traffic Safety Administration; Jacqueline Coolman and Larry Cruse, Government Documents Department, Central Library, University of California, San Diego, for extremely skilled and painstaking assistance; Gail Hall, Wanda Kownacki, Richard Murphy, Allen Mar, Kenneth Chyten, Karen Theobald, Nancy Fackler, and Cathy Mallinger for helping to collect and analyze large quantities of data; Commander Robert Bieber, Analysis Section, California Highway Patrol, for information on procedures for tabulating California motor vehicle fatality statistics; T. J. Donnelly, Supervisor of Communications, Audit Bureau of Circulations, for information on procedures for determining and auditing newspaper circulation figures; and my colleagues Bennett Berger, Fred Davis, Murray Davis, Joseph Gusfield, Gary Marx, Chandra Mukerji, and Jacqueline Wiseman for their comments and criticisms while this research was in progress.

[2] A small part of this paper has been briefly summarized in a two-page report in *Science* (Phillips 1977*a*). The terms "imitate," "suggest," and "model" will be defined throughout this paper as they are in *Webster's Unabridged Dictionary:* (1) imitate: to follow as a pattern, model, or example; (2) suggest: to mention as a possibility, to offer as an idea for consideration; (3) model: a person or thing regarded as worthy of imitation. Although these concepts have been neglected in contemporary sociology, this is of course not the case in classical sociological theory (see, e.g., Tarde 1903; Le Bon 1895; Mead 1934).

of the widespread interest in suggestion and modeling displayed outside sociology. Each year the U.S. advertising industry spends more than $23 billion on the working assumption that commercial behavior is influenced by suggestion, modeling, and imitation. (For information on the amount spent on advertising, see Sandage and Fryburger [1975, p. 10]; for reviews documenting the effects of advertising, see Dirksen and Kroeger [1973]; Michman and Jugenheimer [1976]; Sandage and Fryburger [1975]; for general reviews of factors affecting consumer behavior, see Engel, Kollat, and Blackwell [1973]; Cohen [1972]; Britt [1970]; Jacoby [1976].) The literature on these topics is estimated to include at least 5,000 articles and books (Jacoby 1976). The effect of imitation, suggestion, and modeling has also been extensively documented in a noncommercial context—most notably by psychologists studying the effects of the mass media (especially television) on violence (Klapper 1960; Berkowitz 1962; Flanders 1968; Goranson 1969; Singer 1971; Bandura 1973; Comstock and Fisher 1975; Comstock 1975; Comstock and Lindsey 1975; Liebert and Schwartzberg 1977). To date, there have been more than 2,300 studies on this topic, the most widely known being the Surgeon General's report on the impact of televised violence (1972).

Much of the literature on this topic consists of laboratory-based studies of children or youths. On the basis of these studies, many investigators agree that certain types of television and film programs help to stimulate violence in children in the laboratory (see references above). But investigators disagree strongly on whether these laboratory studies imply that the mass media have any serious effect on adult violence in the real world.[3]

The purpose of this paper is to provide sociological evidence indicating a previously unsuspected, real-world effect of suggestion on one kind of adult violence. In view of the results to be presented, it may be valuable to consider whether the concepts of suggestion, modeling, and imitation can be usefully incorporated into sociological theory. Does the use of these concepts sensitize us to theoretical issues previously neglected and provide us with answers to problems previously unsolved? This question will be addressed after presentation of the data.

DESCRIPTION OF THE STUDY

The study to be presented can be best understood in the light of an earlier paper (Phillips 1974). In that paper I showed that (1) the national level of suicides increases significantly for a brief period after a suicide story is

[3] For a discussion of this issue, see Comstock (1975, pp. 36–40). Of course, the preceding discussion should not be taken to imply that social psychological laboratory studies are not valuable, only that the relatively precise causal analysis made possible by such studies needs to be supplemented by other types of study which allow one to determine whether laboratory studies can be generalized to the "real world."

publicized by the newspapers (for example, after Marilyn Monroe committed suicide, the U.S. suicide rate temporarily increased by 12%); (2) this increase in suicides occurs only after the suicide story is published; (3) the more publicity given to the suicide story, the more the national level of suicides increases; and (4) the increase in suicides occurs primarily in the geographic areas where the suicide story is publicized. The increase in suicides after a publicized suicide story remains after one corrects for the effects of random, seasonal, and yearly fluctuations on the suicide level. After testing several alternative explanations of these findings, I concluded that suicides increase after a publicized suicide story because some persons are imitating the publicized suicide. The phenomenon thus discovered was labeled "the Werther effect," after Goethe's literary hero who committed suicide, and was subsequently imitated.

As long as one can assume that the Werther effect is an isolated sociological phenomenon, one need not consider whether concepts like suggestion should be added to the body of sociological theory. If, however, the effect turns out to be not an isolated phenomenon, but part of a general category of similar phenomena, it may be useful for sociologists to pay more serious attention to concepts like suggestion.

In this paper I will address a host of questions left unanswered in the earlier study. In addressing these questions, I will show (among other things) that (1) suggestion affects behavior which was previously thought to be unaffected by it; (2) different types of publicized stories exert different types of effect; (3) publicized stories affect some types of people more than others.

THE EFFECT OF SUGGESTION ON DRIVERS IN MOTOR VEHICLE ACCIDENTS

Motor vehicle accidents are the fifth leading cause of death in our society (U.S. Public Health Service 1974). Although the evidence is far from conclusive, many researchers have conjectured that some motor vehicle accidents have a suicidal component.[4] If this is so, motor vehicle accidents

[4] For example, see Litman and Tabachnick 1967; Porterfield 1960; MacDonald 1964; Tabachnick et al. 1966, 1973; Preston 1964; Preston and Harris 1965; Conger et al. 1957, 1959; Brown and Birdie 1960; Tillman and Hobbs 1949; Selzer 1969; Finch and Smith 1970; Selzer and Payne 1962; Crancer and Quiring 1970. These investigations are based largely on psychiatric or psychological studies of small, nonrandom samples restricted to very small geographic areas. In a typical study of this type, the researcher seeks to find a similarity between the psychological profile of the auto accident victim and the psychological profile of the suicide. In some studies, a similarity is found (e.g., Tabachnick et al. 1966); in others, sometimes by the same investigators, no similarity is found (Tabachnick et al. 1973; Preston 1964). For a more complete review of the contradictions and methodological problems inherent in this literature, see Phillips (1977*a*).

should behave just like suicides; that is, they should increase very markedly just after publicized stories about suicides. I will now show that motor vehicle accidents do indeed behave in this fashion.

Method of Analysis

The mortality data to be examined consist of tables of daily motor vehicle fatalities. Four states publish such information (California, Michigan, Texas, and South Carolina),[5] but only the largest of the four (California) will be examined here. In 1970, the midpoint of the period to be studied, California accounted for 9.81% of all U.S. motor vehicle fatalities (U.S. Public Health Service 1974). The remaining three states will be examined in a subsequent paper.

An exhaustive list of front-page suicide stories was generated from the reference libraries of the *Los Angeles Times* and the *San Francisco Chronicle*.[6] These are, respectively, the largest newspapers in the Los Angeles and the San Francisco Standard Metropolitan Statistical Areas, the two largest Standard Metropolitan Statistical Areas in the state. Three additional large California papers will also be investigated below.

If front-page California suicide stories stimulate a rise in California motor vehicle fatalities, then this increase can be detected by a technique adapted from earlier work (Phillips and Feldman 1973; Phillips 1974; Phillips 1977*a*). The use of this technique can be illustrated by the case of Yukio Mishima, a Japanese author who killed himself amid great international publicity on Tuesday, November 24, 1970. In the "experimental period" consisting of the week after Mishima's death (November 24–November 30), there were 117 motor vehicle fatalities in California.[7] One can determine whether this is an abnormally large number of deaths[8] in the experimental

[5] Personal communication from Grace Hazzard and Reuben Harding, Statistical Information Staff, National Highway Traffic Safety Administration.

[6] In keeping with previous practice (Phillips 1974), only front-page (not inside-page) newspaper stories have been examined. One might also wish to examine the effect of suicide stories publicized on television. Unfortunately, it is very difficult to compile a systematic list of stories publicized by the 52 television stations in California in 1970 (Television Digest 1970) because these stations do not in general keep an index of the stories they have covered through the years

[7] At the beginning of this study the "experimental period" was arbitrarily set at one week instead of, say, a month. This choice was made because it was assumed that the effect (if any) of a suicide story on motor vehicle deaths would last about a week. In general, for all the suicide stories studied, the experimental week starts with the earliest date on which publicity might be devoted to the suicide story. Consequently, the "experimental week" starts on the day of the suicide, not on the day when the suicide story appears in the newspaper (which is often a day after the suicide occurs). Following the above procedure allows one to detect the early effects of television or radio publicity which may be devoted to a suicide before it is described in the newspapers.

[8] Information on the daily number of California motor vehicle fatalities is obtained from California Highway Patrol, *Annual Report of Fatal and Injury Motor Vehicle*

period by comparing this number (117) with the number of deaths occurring in "control periods" in other years. These control periods are matched with the experimental period in several ways:

1. Month of the year. Since the experimental period falls in November, the control periods are chosen so as to fall in November also. Matching control periods and experimental periods with respect to month corrects for the effect of seasonal fluctuations on motor vehicle fatalities.

2. Day of the week. Motor vehicle fatalities vary markedly by day of the week;[9] thus, it is advisable to match control periods and experimental periods with respect to days of the week. In the present example, this means choosing one-week control periods which fall at the end of November *and* begin on Tuesday, just as the experimental period does. Matching control periods and experimental periods with respect to day of the week corrects for the effect of weekday fluctuations on motor vehicle fatalities.

3. Presence or absence of holiday weekends. Motor vehicle fatalities increase markedly during holiday weekends; hence, if the experimental period contains a holiday weekend, the control periods should also. Conversely, if the experimental period does not contain a holiday weekend, the control periods should not either. Matching control and experimental periods with respect to holiday weekends corrects for the effects of holiday weekends on motor vehicle fatalities.

For the period under analysis in this paper, 1966–73,[10] four control periods contain no publicized suicides and can be matched with the experimental period, Tuesday, November 24–Monday, November 30, 1970, in the ways described above.

The number of motor vehicle fatalities in each of these four control periods was calculated, and a regression line was fitted to these data. From this regression line, one would expect 98.88 motor vehicle fatalities in the experimental period, November 24–30, 1970, under the null hypothesis that publicized suicides have no effect on motor vehicle fatalities. There was a rise in motor vehicle deaths in the week after Mishima killed himself. The observed number of fatalities in this period (117) is greater than the number expected (98.88) ($P < .05$, one-tailed t-test).

Traffic Accidents (yearly issues). For the week after the publicized suicide, one might wish to examine fatality rates rather than the number of fatalities Unfortunately, it is not possible to calculate weekly fatality rates because the weekly population figures needed for these calculations do not exist.

9 This is immediately apparent from inspection of daily fatality figures, as is the marked increase during holiday weekends mentioned in the same paragraph.

10 The period of analysis begins in 1966 because that is the earliest year for which the California Highway Patrol published daily motor vehicle accident fatality statistics. The period of analysis ends in 1973 because that was the last year before the Arab oil embargo drastically changed patterns of gasoline consumption and patterns of motor vehicle traffic.

In general, the procedure just described was used to estimate the effect of all front-page suicide stories in California for which suitable control periods were available.[11] Some suicide stories could not be investigated because, for example, their control periods contained suicide stories and hence could not be used.

Results of Analysis

Table 1 gives the number of California motor vehicle fatalities observed after each front-page suicide story, the number expected under the null hy-

TABLE 1

RISE IN MOTOR VEHICLE FATALITIES AFTER PUBLICIZED SUICIDES, CALIFORNIA 1966–73

Identity of Publicized Suicide	Date of Publicized Suicide	Observed No of Motor Vehicle Deaths in the Week after the Suicide	Expected No of Motor Vehicle Deaths in the Week after the Suicide†	% Rise in Motor Vehicle Deaths 100 ([Observed − Expected]/Expected)
A. Korbel, winemaker.	4-21-1966	96	89 67	7 06
S. Youngren, union leader	5-17-1966	79	80 40	−1 74
Lo Jui-Ching, army leader	1-19-1967	81	80 92	10
J. Hughes, businessman*	2-23-1967			
E. Joe, student*	2-23-1967	102	78 90	29 28
S. Abshear, student*	2-25-1967			
A. Amer, Egyptian general....	9-14-1967	99	101 20	−2.17
M. Berg, L.A. policeman	9-23-1967	105	100 70	4.27
V. Janko, Czech general	3-14-1968	88	84 25	4 45
F. Chegwin, mass murderer	8-8-1968	100	89 40	11 86
H. Luedke, NATO admiral*	11-10-1968	120	89 14	34 62
V. Latham, housewife*	11-10-1968			
J. Palach, Czech student	1-16-1969	102	78 08	30 64
Y. Mishima, author....	11-24-1970	117	98 88	18.33
J. Mattison, prisoner	2-25-1971	90	85 61	5 13
B. Pollack, orchestra leader	6-7-1971	82	102 70	−20 16
G. Giffe, hijacker	10-4-1971	104	97 50	6 67
G. Logan, mass murderer	11-26-1971	89	82 57	7 79
J. Van Praag, psychologist	3-7-1972	103	75 67	36 12
M. Oufkir, defense minister	8-17-1972	101	83.89	20 40
M. Brody, millionaire....	1-26-1973	65	81 00	−19 75
E. Brudno, ex-P.O.W.	6-3-1973	98	102 14	−4 05
W. Inge, playwright ..	6-10-1973	100	88 03	13 60

* Suicides occurring within one week of each other and therefore treated as one story

† See text for data sources and for methods of calculating expected number of deaths.

11 When the number of deaths expected in the experimental period could be estimated from a regression line interpolated between control periods, a minimum of two control periods was required. When the expected number of deaths had to be extrapolated from control periods, a more conservative procedure was followed: (1) A minimum of three control periods was required. (2) A suicide story was not examined if one needed to extrapolate the regression line for more than one year. Researchers have often recommended more conservative estimation procedures when estimates must be based on extrapolated data.

thesis that suicide stories have no effect on motor vehicle fatalities, and ɛ difference between the observed and the expected numbers of fatalities. ı the average, the number of motor vehicle fatalities increased by 9.12% the week after each suicide story ($P = .011$, t-test, one-tail).[12]

ming of Motor Vehicle Fatalities in Relation to the Publicized Suicide

the observed rise in motor vehicle fatalities is caused by publicized suicide ɔries, then this rise should occur only after each story appears. Table 2 ables us to examine the day-by-day fluctuation of motor vehicle fatalities r a two-week period, beginning two days before the publicized suicide and ıding 11 days afterward.[13] It is evident that motor vehicle fatalities do not se immediately before the publicized suicide and do rise rather steeply

TABLE 2

DAILY FLUCTUATION IN MOTOR VEHICLE FATALITIES FOR A TWO-WEEK PERIOD BEFORE, DURING, AND AFTER PUBLICIZED SUICIDES

No of Days before or after Publicized Suicides	No of Deaths Observed	No of Deaths Expected	Percentage Fluctuation: 100 ([Observed − Expected]/Expected)
days before	237	247.038	−4 06
day before	264	268 714	−1 75
ay of publicized suicide	260	247.716	+4 96
day after	269	254 037	+5 89
days after	313	300 624	+4 12
days after	338	257.440	+31.29
days after	264	244.205	+8.11
days after	244	236 771	+3 05
days after	233	225 477	+3 34
days after	226	236.694	−4 52
days after	304	278 341	+9.22
days after	329	315.091	+4.41
) days after	288	291 293	−1.13
l days after	240	236.073	+1.66

NOTE.—For the method of calculating the number of deaths expected, see text.

[12] The statistical techniques used in this paper are plausible but arbitrary, in the sense ıat other, equally plausible techniques might have been substituted for the ones actu- lly used. The results (table 1) remain statistically significant even if these alternative chniques are used: (1) When these results are analyzed by the Walsh test, $P = .010$, ne-tailed test. (2) The expected number of deaths after a story might plausibly be stimated by averaging the deaths in the control periods rather than by fitting these eaths to a regression line. In this case, auto fatalities rise by an average of 7.15%, $P =$)19, one-tailed t-test; $P = .016$, one-tailed Walsh test. In view of these findings, it is ifficult to conclude that the results in table 1 are an artifact of the statistical technique sed to analyze them.

[13] The techniques described earlier for calculating the expected number of deaths in any iven week are adapted for calculating the expected number of deaths on any given day. 'hese techniques can be illustrated by the case of Mishima's suicide on Tuesday, No-

immediately after the suicide. The most noteworthy feature of this table is the brief but very strong increase in motor vehicle fatalities three days after the publicized suicide, at which time motor vehicle fatalities are 31% (and 80 deaths) higher than expected. The data in table 2 are graphed in figure 1.

I have now shown that motor vehicle accident fatalities increase significantly after publicized suicide stories, and not before. In the next section I will provide additional evidence suggesting a causal link between suicide stories and a rise in motor vehicle fatalities.

Relationship between Amount of Publicity Devoted to Each Suicide Story and Increase in Fatalities after Each Story

If motor vehicle fatalities rise because of suicides publicized by the newspapers, two additional predictions should hold true: (1) an index of the

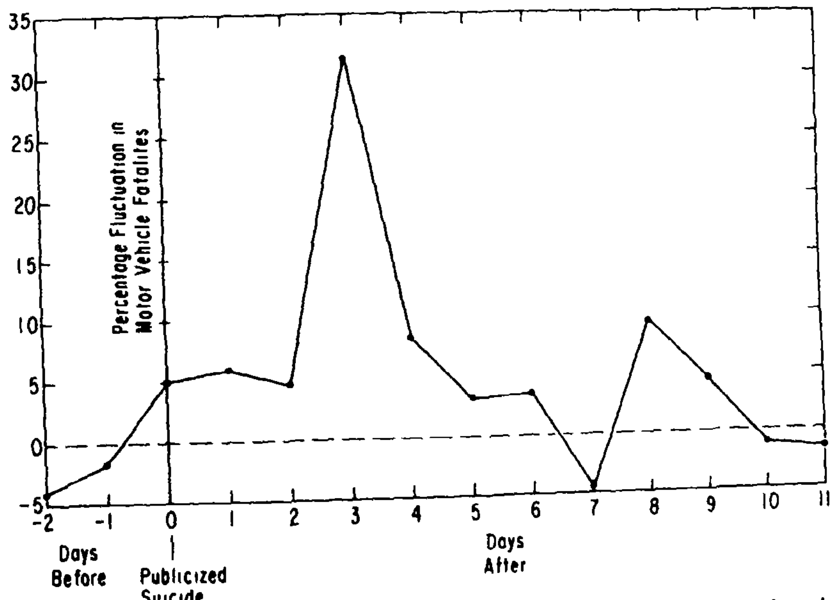


Fig. 1.—Daily fluctuation in motor vehicle accident fatalities for a two-week period before, during, and after publicized suicides.

---

vember 24, 1970. The expected number of deaths on this day is calculated by fitting a regression line to the four control days: Tuesday, November 22, 1966; Tuesday, November 21, 1967; Tuesday, November 26, 1968; and Tuesday, November 21, 1972. A similar regression period is used to calculate the expected number of deaths on the day of each of the publicized suicides. Table 2 indicates that the sum of these 20 expected numbers is 247.716, the total number of deaths expected on the day of the publicized suicides. The remaining expected numbers in table 2 were calculated in a similar fashion.

pothesis that suicide stories have no effect on motor vehicle fatalities, and the difference between the observed and the expected numbers of fatalities. On the average, the number of motor vehicle fatalities increased by 9.12% in the week after each suicide story ($P = .011$, t-test, one-tail).[12]

Timing of Motor Vehicle Fatalities in Relation to the Publicized Suicide

If the observed rise in motor vehicle fatalities is caused by publicized suicide stories, then this rise should occur only after each story appears. Table 2 enables us to examine the day-by-day fluctuation of motor vehicle fatalities for a two-week period, beginning two days before the publicized suicide and ending 11 days afterward.[13] It is evident that motor vehicle fatalities do not rise immediately before the publicized suicide and do rise rather steeply

TABLE 2

DAILY FLUCTUATION IN MOTOR VEHICLE FATALITIES FOR A TWO-WEEK PERIOD BEFORE, DURING, AND AFTER PUBLICIZED SUICIDES

No of Days before or after Publicized Suicides	No. of Deaths Observed	No of Deaths Expected	Percentage Fluctuation: 100 (|Observed − Expected|/Expected)
2 days before	237	247.038	−4.06
1 day before	264	268.714	−1 75
Day of publicized suicide	260	247 716	+4 96
1 day after	269	254 037	+5.89
2 days after	313	300.624	+4.12
3 days after	338	257 440	+31 29
4 days after	264	244 205	+8.11
5 days after	244	236 771	+3 05
6 days after	233	225 477	+3 34
7 days after	226	236.694	−4 52
8 days after	304	278.341	+9.22
9 days after	329	315 091	+4.41
10 days after	288	291.293	−1.13
11 days after	240	236.073	+1.66

NOTE.—For the method of calculating the number of deaths expected, see text.

12 The statistical techniques used in this paper are plausible but arbitrary, in the sense that other, equally plausible techniques might have been substituted for the ones actually used. The results (table 1) remain statistically significant even if these alternative techniques are used: (1) When these results are analyzed by the Walsh test, $P = .010$, one-tailed test. (2) The expected number of deaths after a story might plausibly be estimated by averaging the deaths in the control periods rather than by fitting these deaths to a regression line. In this case, auto fatalities rise by an average of 7.15%, $P = .019$, one-tailed t-test; $P = .016$, one-tailed Walsh test. In view of these findings, it is difficult to conclude that the results in table 1 are an artifact of the statistical technique used to analyze them.

13 The techniques described earlier for calculating the expected number of deaths in any given week are adapted for calculating the expected number of deaths on any given day. These techniques can be illustrated by the case of Mishima's suicide on Tuesday, No-

immediately after the suicide. The most noteworthy feature of this table is the brief but very strong increase in motor vehicle fatalities three days after the publicized suicide, at which time motor vehicle fatalities are 31% (and 80 deaths) higher than expected. The data in table 2 are graphed in figure 1.

I have now shown that motor vehicle accident fatalities increase significantly after publicized suicide stories, and not before. In the next section I will provide additional evidence suggesting a causal link between suicide stories and a rise in motor vehicle fatalities.

Relationship between Amount of Publicity Devoted to Each Suicide Story and Increase in Fatalities after Each Story

If motor vehicle fatalities rise because of suicides publicized by the newspapers, two additional predictions should hold true: (1) an index of the

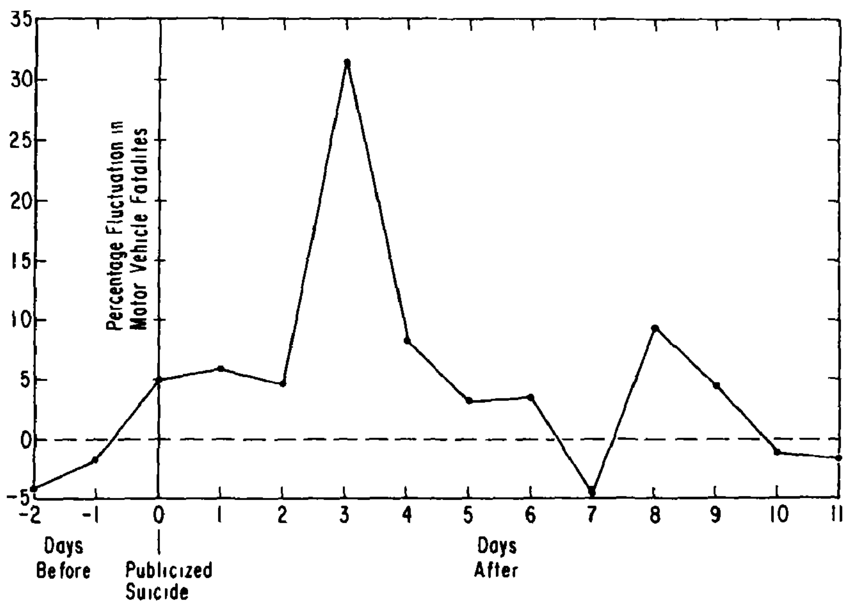


Fig. 1.—Daily fluctuation in motor vehicle accident fatalities for a two-week period before, during, and after publicized suicides.

vember 24, 1970. The expected number of deaths on this day is calculated by fitting a regression line to the four control days: Tuesday, November 22, 1966; Tuesday, November 21, 1967; Tuesday, November 26, 1968; and Tuesday, November 21, 1972. A similar regression period is used to calculate the expected number of deaths on the day of each of the publicized suicides. Table 2 indicates that the sum of these 20 expected numbers is 247.716, the total number of deaths expected on the day of the publicized suicides. The remaining expected numbers in table 2 were calculated in a similar fashion.

newspaper publicity devoted to a story should correlate positively with the rise in motor vehicle fatalities after that story; (2) the more accurate this index of publicity, the more highly it should correlate with the rise in motor vehicle fatalities.

A much-publicized story is one covered by many newspapers with a large combined circulation; a little-publicized story is one covered by few newspapers with a small combined circulation. Ideally, one should examine all 150 or so California daily newspapers to determine the total circulation achieved by a particular story. In practice, the most one can readily achieve is an examination of the five largest papers, which accounted for 41% of the total 1970 daily newspaper circulation in California.[14]

The publicity devoted to a story by these five newspapers (table 3, col. 2) can be used as an index of the total amount of publicity devoted by all newspapers. For any given story, the value of this index was calculated from

TABLE 3

RELATIONSHIP BETWEEN THE AMOUNT OF PUBLICITY DEVOTED BY THE FIVE BIGGEST CALIFORNIA NEWSPAPERS TO EACH SUICIDE STORY AND THE RISE IN MOTOR VEHICLE FATALITIES AFTER EACH STORY

Name of Publicized Suicide*	Total Circulation of Newspapers Covering the Story	Rise in Motor Vehicle Fatalities in the Week after the Story (%)
Palach	2,627,084	+30.64
Logan	1,858,095	+7.79
Lo Jui-Ching	1,799,279	+.10
Van Praag	1,673,972	+36.12
Mishima	1,641,766	+18.33
Luedke and Latham	1,578,018	+34.62
Hughes, Joe, and Abshear	1,414,326	+29.28
Giffe	1,376,928	+6.67
Janko	1,348,430	+4.45
Oufkir	1,022,359	+20.40
Inge	1,004,908	+13.60
Pollack	966,293	−20.16
Brudno	951,263	−4.05
Chegwin	856,621	+11.86
Amer	847,869	−2.17
Berg	847,869	+4.27
Korbel	750,000	+7.06
Youngren	750,000	−1.74
Mattison	478,704	+5.13
Brody	458,163	−19.75

* To facilitate analysis, suicide stories have been listed in the order of the amount of publicity received by each story.

14 Information on newspaper circulation figures and on the procedures for ensuring that these figures are accurate is provided by *Ayer Directory of Publications* (yearly volumes) and by the Audit Bureau of Circulations, Chicago.

$$\sum_{i=1}^{5} x_i y_i ,$$

where x_i is the circulation of newspaper i at the time of the story, and y_i is the number of days that newspaper carried the story on page 1. This five-newspaper publicity index correlates positively ($r = .59$, $P < .005$) with the change in motor vehicle fatalities after each story (table 3, col. 3). On the average, fatalities increased by 18.84% after stories receiving more than the median amount of publicity and decreased by a statistically insignificant amount, 0.60%, after stories receiving less than the median amount of circulation.

In table 3, a five-newspaper circulation index was used. In general, the more newspapers included in a circulation index, the more highly it should correlate with the rise in motor vehicle fatalities after publicized suicide stories. In all, five indexes have been constructed. These range from index 1 (constructed from circulation data for the single largest California newspaper) to index 5 (constructed from circulation data for the five largest California newspapers). Each index is correlated with the fluctuation of motor vehicle fatalities as follows: index 1, $r = .196$; index 2, $r = .300$; index 3, $r = .478$; index 4, $r = .576$; and index 5, $r = .594$. As predicted, the more newspapers included in the index of publicity, the higher the correlation with motor vehicle fatalities. These five indexes would be ranked in the predicted order by chance .0083 (1/120) of the time.

We have now seen that motor vehicle accidents increase sharply just after publicized suicide stories. We have also seen that the more publicity devoted to a suicide story, the more motor vehicle fatalities rise thereafter. The more precisely one measures publicity, the more strongly it correlates with the rise in motor vehicle fatalities.

Some Possible Alternative Explanations of the First Set of Findings

The findings I have presented above are consistent with the hypothesis that publicized suicide stories help to stimulate a brief increase in motor vehicle fatalities. The results cannot be ascribed to the effect of weekday or monthly fluctuations in motor vehicle fatalities, to holiday weekends, or to yearly linear trends, because these effects were corrected for in the selection and treatment of the control periods with which the experimental periods are compared. Furthermore, because the results are statistically significant, they cannot plausibly be ascribed to chance fluctuations in the data.

The hypothesis that a prior change in social conditions caused both the publicized suicide and the subsequent rise in motor vehicle fatalities is implausible for two reasons: (1) if such conditions created both a wave of motor vehicle fatalities and the front-page suicide, the suicide would be

expected to occur during the wave of fatalities rather than before it; (2) the prior-conditions explanation implies that there is no link between the characteristics of front-page news stories and the increase in motor vehicle fatalities. If no such link exists, it is difficult to explain the observed association between the publicity given to a suicide and the rise in fatality levels thereafter. Additional, minor alternative explanations have been examined elsewhere (Phillips 1977*b*).

The best available explanation of the rise in fatalities is that it is caused by the suggestive effect of suicide stories. The precise nature of this effect will be much better understood after one examines the characteristics of the accidents and the characteristics of the drivers dying in those accidents.

DETAILED ANALYSIS OF TYPES OF ACCIDENTS OCCURRING JUST AFTER SUICIDE STORIES

In this section I will attempt to answer questions like the following: Do stories about suicide trigger single-car accidents rather than multiple-car accidents? Do stories about suicide combined with murder trigger multiple-car accidents (this being a method by which a driver can kill others as well as himself)? Is the driver who dies after a suicide story similar to the person described in that suicide story? In seeking to answer questions like these one gains insight into the social and psychological processes producing the observed increase in motor vehicle fatalities. In addition one gains a more general understanding of the processes of imitation, suggestion, and modeling.

Selection of Accidents to Be Examined

When one attempts to determine the characteristics of accidents occurring after suicide stories, it is inefficient to examine all accidents that occurred throughout the state of California. One would prefer, instead, to include only those adult decedents who were most likely to have been exposed to the suicide stories under study. These stories were all publicized in the Los Angeles or San Francisco Standard Metropolitan Statistical Area and were not necessarily publicized elsewhere in California. Consequently, I will now restrict attention to adult Californians who resided and/or died in the Los Angeles or San Francisco Standard Metropolitan Statistical Area in 1969–73. Mexican-Americans will be excluded from the analysis because surveys show that this ethnic group is much less likely than others to read the largest California newspapers.[15]

[15] The published tabulations used in the first part of this paper pertain to all Californians for the period from 1966 to 1973. They are not subclassified by age, race, geographic location in California, type of accident, etc. With the computerized death certificates used in the second part, these subclassifications are possible but, in some cases, only for 1969 and thereafter. Consequently, the analysis in this part of the study must

In examining accidents following suicide stories, one would prefer to focus attention on the "suicidal" accidents which were triggered by the suicide story rather than the "conventional" accidents which would have occurred anyway, even in the absence of a suicide story. Unfortunately, the available data do not allow us to separate suicidal from conventional accidents; the best we can do is examine all accidents in the time period which contains the largest possible ratio of suicidal accidents to conventional ones.

The method for selecting this time period can best be understood if we refer again to figure 1, which shows the fluctuation of motor vehicle fatalities before and after publicized suicide stories. The dotted line in the graph indicates the level of motor vehicle fatalities expected if fatalities were unaffected by suicide stories. Accidents indicated below the dotted line are presumably conventional motor vehicle fatalities which would be expected to occur even in the absence of suicide stories. The excess fatalities indicated above the dotted line of the graph presumably include some suicidal accidents which were triggered by the publicized suicide stories. It is evident from this graph that the largest number of excess fatalities occurs on the third day after each publicized suicide. Hence, this day should contain the largest proportion of suicidal automobile fatalities. Thus, I have chosen the third day after each story as the experimental period in this part of the study.

The control period to be used in this part of the study consists of the seventh day after the experimental period and the seventh day before the experimental period. Thus, the control period always falls on the same day of the week as the experimental period and, in general, control periods and experimental periods will also be matched with respect to month and year.[16] This method of matching experimental and control periods allows us to correct for the effect of weekday, monthly, and yearly fluctuations in the characteristics of motor vehicle fatalities. Therefore, any differences found

be restricted to 1969–73 instead of the longer period used earlier In general it should be stressed that the data base for this part of the paper is a subset of that employed above The following definitions have been used in this part of the study: A Californian is someone who resided and died in California. An adult is someone aged 15 or over. A Mexican-American is defined as an American who was born in Mexico One might prefer to broaden this definition to include (say) all Spanish-speaking Americans, but unfortunately the information needed to apply a broader definition is not available in the computerized death certificates used in this study The *Los Angeles Times* readership study showed that only 13.5% of adult Mexican-Americans in Los Angeles read the weekday *Los Angeles Times;* in contrast, 36.15% of all other adults read the *Los Angeles Times.*

16 The method for choosing control periods in the first part of this study can not be used in this part of the study. The method used in the first part of the study is feasible when a relatively large span of years is examined (1966–73) but not for the shorter span of years covered in this current section of the paper.

between the two periods will not be attributable to weekday, monthly, or yearly fluctuations in the data. In all, 169 deaths will be examined in this part of the study—62 deaths in the experimental period and 107 deaths in the control period. Because the control period consists of twice as many days as the experimental period, it is expected to have more deaths.

I will now examine the characteristics of "experimental group" accidents (which follow suicide stories) and the characteristics of "control group" accidents (which occur at other times).

Are Single-Vehicle Accidents Abnormally Frequent Just after Suicide Stories?

Researchers have often supposed that suicidal accidents are more likely to be single-vehicle accidents than to be other types.[17] If this is so, the ratio of single-vehicle accident deaths to other types should be larger in the experimental period than in the control period. This prediction is confirmed by the data in table 4: the ratio of deaths from single-vehicle accidents to deaths from other types is 29:33 = .88 in the experimental period, while the ratio is much smaller in the control period (32:75 = .43). The difference between these two ratios is significant at .0213 (hypergeometric, one-tail); hence deaths from single-vehicle accidents increase more than other types just after a publicized suicide story.[18]

TABLE 4

RELATIVE FREQUENCY OF DEATHS FROM SINGLE-VEHICLE ACCIDENTS JUST AFTER SUICIDE STORIES ARE PUBLICIZED—EXPERIMENTAL PERIOD AND OTHER TIMES (Control Period)

	Deaths from Single-Vehicle Accidents	Deaths from Other Motor Vehicle Accidents
Experimental period . .	29	33
Control period . .	32	75

NOTE.—"Single-vehicle accidents" are defined in n. 17.

17 The literature on this topic was reviewed earlier in this paper (in n. 4 above). A single-vehicle accident is defined here as an accident coded as E815 or E816 in the Eighth International Classification of Causes of Death; E815 and E816 comprise all motor vehicle collision traffic accidents not involving a train, streetcar, another motor vehicle, another vehicle, or a pedestrian.

18 This application of the hypergeometric is discussed by Hoel (1965) and even more specifically by Mosteller, Rourke, and Thomas (1970).

Do the Drivers in Single-Vehicle Accidents Identify with the Person in the Suicide Story?

If the drivers in single-vehicle accidents have identified with the person in a suicide story, then suicide stories about young persons should tend to induce single-vehicle crashes with young drivers. Conversely, suicide stories about older persons should induce single-vehicle crashes with older drivers. These predictions are supported by the available data: there is a positive correlation ($r = .4634$, $P = .023$, t-test, one-tailed, 16 df) between the age of the publicized suicide and the age of drivers in single-vehicle crashes which occur just after the publicized suicide.[19]

The findings presented above indicate that suicide stories may lead to relatively many single-vehicle crashes but relatively few multiple-vehicle crashes. Hence, it is probable that only a small proportion of multiple-vehicle crashes are caused by suicidal persons who have identified with the publicized suicide. Thus, one would not expect to find a strong correlation between the age of the publicized suicide and the age of drivers in *multiple*-car crashes which occur just after the publicized suicide. As predicted, r is small ($= .0655$, $P = .416$, t-test, one-tailed, 10 df). A similarly small, insignificant correlation is found when one compares the age of the publicized suicide with the age of *non*drivers who die just after the suicide story ($r = .2109$, $P = .146$, t-test, one-tailed, 25 df).[20]

Do Different Types of Suicide Stories Trigger Different Types of Motor Vehicle Accidents?

Two major types of suicide story can be distinguished: (1) pure suicide stories, in which a person kills himself and no one else, and (2) murder-suicide stories, in which a person kills others as well as himself. One might expect that pure suicide stories would tend to trigger single-car accidents in which the driver dies, while murder-suicide stories would tend to induce multiple-car accidents in which passengers die. Table 5 shows the relative frequency with which deaths from different types of accident occur after each type of suicide story. Although the numbers in this table are quite

[19] Some of the assumptions necessary for testing the statistical significance of this correlation may not hold. Thus, one might prefer an alternative statistic (e g., the Spearman rank-correlation coefficient) which requires less stringent assumptions. The Spearman rank correlation $= .444$ (corrected for ties). This figure is significant at $< .05$, one-tailed. It would be interesting to determine whether the race and sex of the publicized suicide are correlated with the race and sex of drivers dying just after the suicide story. Unfortunately, one cannot make this determination because publicized suicides hardly vary with respect to race and sex—almost all are white males.

[20] As one would expect, there is no significant correlation between the age of the publicized suicide and the age of persons dying in the control period.

TABLE 5

CHARACTERISTICS OF MOTOR VEHICLE ACCIDENT DEATHS FOLLOWING SUICIDE AND MURDER-SUICIDE STORIES

	Driver Deaths from Single-Vehicle Accidents (1)	Driver Deaths from Multiple-Vehicle Accidents (2)	Passenger Deaths from Single-Vehicle Accidents (3)	Passenger Deaths from Multiple-Vehicle Accidents (4)	Pedestrian Deaths from Motor Vehicle Accidents (5)	Deaths from Other (Miscellaneous) Accidents (6)
Suicide stories .	15	12	5	2	5	5
Murder-suicide stories...... .	3	5	4	6	0	0

NOTE.—The deaths in each column correspond to the following categories in the Eighth International Classification of Causes of Death 8152, 8162, 8150, 8160 (col 1), 8102, 8112, 8122, 8132, 8100, 8110, 8120, 8103 (col. 2); 8151, 8161, 8153, 8163 (col. 3), 8101, 8111, 8121, 8131, 8103, 8113, 8123, 8133 (col 4), 8147 (col. 5), 8128, 8129, 8169, 8181 (col. 6). The majority of cases in col. 6 involve residual categories 8129 and 8169. In these accidents the death certificate fails to specify the nature of the person killed (e.g , a driver, passenger, etc).

small, the differences among them are so extreme that the results in table 5 are statistically significant. Table 5 shows that accident deaths following pure suicide stories are significantly different from accident deaths following murder-suicide stories ($\chi^2 = 14.67$, $P = .012$, 5 df). Multiple-vehicle accidents involving passenger deaths tend to be relatively frequent after murder-suicide stories and quite rare after pure suicide stories. Conversely, single-vehicle accidents involving driver deaths are very frequent after pure suicide stories and very rare after murder-suicide stories. These results are consistent with the notion that a pure suicide story stimulates some persons to commit suicide (but not murder), while murder-suicide stories stimulate some persons to commit murder and suicide.[21] (See also Phillips 1978 for further evidence on the effect of murder-suicide stories.)

Do the Crashes after Suicide Stories Occur Mainly in the Area Where the Suicide Stories Are Publicized?

If publicized suicide stories help to produce an increase in motor vehicle fatalities, the increase in fatalities should occur particularly in the places where the suicide story is publicized. This prediction will be tested after some terms are defined.

[21] The apparent overabundance of pedestrian deaths after pure suicide stories but not after murder-suicide stories may be explainable by the following speculative line of reasoning. A pedestrian can imitate a publicized suicide by stepping in front of a car and killing himself. But he cannot imitate a murder-suicide by stepping in front of a car, because he cannot commit murder by this action, only suicide. The results in table 5 must be interpreted with caution because the death certificate data do not indicate the number of persons dying in a given accident. Because of these limitations in the data, one cannot determine (for example) whether drivers dying immediately after a pure suicide story were alone at the time of their accidents.

For each suicide story, one can divide the Los Angeles and San Francisco region into two parts: (1) a "publicity area" consisting of the city or cities known to be receiving publicity on the story, and (2) a "nonpublicity area" consisting of the city not known to be receiving publicity on the story. Thus, for example, a story in the present study may be publicized in Los Angeles and not in San Francisco, in San Francisco and not in Los Angeles, or in both cities at once. For each story s, four statistics can be calculated: (1) $p(\text{exp})_s$, the number of deaths in the publicity area in the experimental period; (2) $p(\text{cont})_s$, the number of deaths in the publicity area in the control period; (3) $\bar{p}(\text{exp})_s$, the number of deaths in the nonpublicity area in the experimental period; and (4) $\bar{p}(\text{cont})_s$, the number of deaths in the nonpublicity area in the control period. Each of these four statistics can then be summed over all stories. For example, one can calculate $\Sigma p(\text{exp})_s$, the total number of deaths in the publicity area in the experimental period for all stories combined.

If motor vehicle fatalities increase mainly in the area where the suicide story is publicized, the ratio of deaths in the publicity area to deaths in the nonpublicity area should be greater for the experimental period than it is for the control period. This prediction is supported by the data in table 6. The two ratios in this table (41:21 = 1.95) and (48:59 = .81) are significantly different at .0058 (hypergeometric, one-tail). In other words, after a suicide story motor vehicle fatalities are abnormally frequent mainly in the area where the suicide story is publicized.

Are Crashes Just after Suicide Stories Unusually Lethal?

I will now examine one final difference between the conventional accidents in the control period and the suicidal accidents in the experimental period. In a suicidal accident the driver is consciously or unconsciously trying to die; hence, just before the crash the driver may well have his foot on the

TABLE 6

NUMBER OF MOTOR VEHICLE FATALITIES IN THE PUBLICITY AREA AND THE NONPUBLICITY AREA DURING THE EXPERIMENTAL PERIOD AND DURING THE CONTROL PERIOD

	Publicity Area	Nonpublicity Area
Experimental period	41	21
Control period	48	59

NOTE.—For definition of the terms "publicity area" and "nonpublicity area," see text.

accelerator, not the brake. The victims[22] of such a suicidal crash should die very soon after the crash occurs, because the driver means the crash to be lethal. On the other hand, the victims of a conventional crash may well take longer to die after the crash occurs, because the driver in a conventional crash does not mean the crash to be lethal. In fact, the driver is trying to avoid the crash: he may well have his foot on the brake, not on the accelerator.

If this argument is correct, the average time between crash and death should be very short for the victims of the suicidal crashes in the experimental period. Conversely, the average time between crash and death should be relatively longer for the victims of the conventional crashes in the control period. These predictions are supported by the data in table 7. For the control period, the average time between crash and death, M_c, is 4.131 days; the average time for the experimental group, M_e, is only 1.016 days. The difference between these two figures, $M_c - M_e$, is significant at .0324 (one-

TABLE 7

NUMBER OF DAYS BETWEEN CRASH AND DEATH, JUST AFTER SUICIDE STORY (Experimental Period) AND AT OTHER TIMES (Control Period)

No Days between Crash and Death	No Cases with x Days between Crash and Death: For the Experimental Group	For the Control Group
0	51	79
1	5	6
2	1	0
3	0	2
4	0	1
5	0	2
6	1	4
7	0	2
8	1	1
9	1	1
10	1	0
11+	1	9

NOTE —For convenience, the data are displayed in grouped form, but the averages discussed in the text (1.016 days for the experimental group and 4 131 days for the control group) were calculated from ungrouped data

[22] In a crash involving a suicidal driver who is accompanied by passengers, the interval between crash and death should be very short for the driver *and* for the passengers. Hence, in this analysis I have examined not only driver deaths but nondriver deaths as well.

tailed t-test for difference between means, 122 df).[23] Thus, persons crashing after a suicide story die much more rapidly than persons crashing at other times.

SUMMARY

In this paper I have shown that: (1) motor vehicle accident fatalities increase markedly just after publicized suicide stories; (2) the increase in accident fatalities occurs only after the publicized story; (3) the more publicity given to the story, the greater the increase in motor vehicle fatalities; (4) the increase in fatalities occurs mainly in the area where the suicide story is publicized; (5) single-vehicle crashes increase more than other types just after suicide stories; (6) suicide stories about young persons tend to be followed by single-vehicle crashes involving young drivers; conversely, suicide stories about older persons tend to be followed by single-vehicle crashes involving older drivers; (7) stories about murder and suicide tend to be followed by multiple-vehicle crashes involving passenger deaths, while stories about suicide alone tend to be followed by single-vehicle crashes involving driver deaths. Thus, pure suicide stories and murder-suicide stories seem to trigger different types of crashes.

DISCUSSION

These findings indicate strongly that drivers of motor vehicles are affected by the processes of imitation, suggestion, and modeling. In view of these and previous findings on the impact of suggestion, it may be valuable to consider whether the concepts of suggestion, imitation, and modeling can be usefully incorporated into sociological theory. Does the use of these concepts sensitize us to theoretical issues previously neglected or overlooked? Can

[23] The difference, $M_c - M_e = 3.115$, cannot be ascribed to differences between the age distributions of the control group and the experimental group, because the difference $M_c - M_e$ actually increases from 3.115 to 3.314 after one corrects for age differences in the control and experimental groups. The statistical significance of the difference $M_c - M_e$ was calculated according to formulas 10.13 1–10 13 4, 10.17 1 in Hays (1963) Statistics used in these calculations are: $N_c = 107$ = number of items in the control group; $N_e = 62$ = number of items in the experimental group; estimated $\sigma^2 M_c = 2.616$; estimated $\sigma^2 M_e = .200$; $t = 1.8564$; df = 122 as calculated from formula 10 17.1 in Hays. Some of the assumptions required by the t-test may not be met by the data; hence, one might prefer an alternative statistical test with less stringent assumptions, e.g , the randomization test (Siegel 1956, pp. 152–56). The difference $M_c - M_e$ is significant at .031, randomization test, one-tail. The accidents in the control period and in the experimental period do not appear to differ significantly with respect to the other major variables coded on the computerized death certificate: race, sex, marital status, known involvement with alcohol or drugs.

one use concepts like suggestion to provide partial answers to problems previously unsolved? In considering these questions, I will illustrate the discussion with references to three very different areas of sociological theory: (1) behavioral decision theory; (2) reference group theory; and (3) the theory of anomie and its effects.[24]

Possible Usefulness of the Concepts of Suggestion, Imitation, and Modeling in Behavioral Decision Theory

In this theory, the individual is conceived of as a decision maker who analyzes a problem rationally, calculates the costs and benefits of various solutions, and, on the basis of these calculations, chooses an optimal solution (for reviews and general discussions of this topic, see Raiffa [1970]; Edwards and Tversky [1967]; Williams [1954]; Slovic, Fischoff, and Lichtenstein [1977]).

Decision theorists start by assuming a particular problem and then consider *how the individual assesses possible solutions to that problem.* By proceeding in this way, decision theorists ignore an important prior question: How does the individual choose the *problem* to be solved, before he even begins to concern himself with solutions? Out of the enormous range of problems plaguing any individual, what makes him single out one for attention rather than another?

Concepts like suggestion and modeling may provide a partial solution to this question. One may pay particular attention to a certain problem if that problem is made salient by a person who serves as a model. Thus, after Mrs. Nelson Rockefeller and Mrs. Gerald Ford discovered that they had breast cancer, breast cancer became a more salient problem for women across the United States. Similarly, after a fictional individual in a television commercial becomes aware of "bad breath," this problem may become salient for many viewers. Thus, suggestion and modeling may function as a kind of mediating mechanism, a kind of "mental switch" which leads an individual to become preoccupied with one problem rather than another.[25] After the individual has become concerned with a particular problem, then perhaps he may choose a solution to that problem according to the criteria of decision theory.

24 These three theories were chosen because: (1) each theory has generated a very large literature; (2) each is intended to cover a very wide range of behaviors; and (3) all three seem to overlook some basic questions which are related in some surprising and interesting ways (as will be indicated in the rest of this section).

25 The individual is not necessarily conceived of here as a helpless puppet, twitched one way or another by all-powerful suggestions and models. He may play an active role in the process of imitating, for example by seeking out favored models he wishes to imitate. This view of the individual as an active model seeker is consistent with a great deal of research on the effects of advertising (see, e.g., Sandage and Fryburger 1975).

A second, basic question ignored by decision theorists is: Why does an individual sometimes approach a problem in a deliberate, reasoning, and logical mode, while at other times he may approach the same problem in an impulsive, unreasoned, and illogical mode? In short, why does an individual behave according to the criteria of decision theory at some times and not at others?

Once again, suggestion, imitation, and modeling may provide a partial answer. An individual may adopt a decision-theoretic mode if that mode is displayed by his parents, teachers, or by other salient models. On the other hand, a non-decision-theoretic mode may be adopted if this alternative mode is made salient. Thus, for example, a child may solve a problem in an impulsive, unreasoned, and violent fashion if this mode is displayed on television. An adult may lash out violently and impulsively, killing others and himself, if this non-decision-theoretic mode is displayed by the media.

Here, as before, the processes of suggestion and imitation may serve as mediating mechanisms which help to switch an individual from one choice to another—in this case from a decision-theoretic mode to another mode of problem solving.

Possible Usefulness of the Concepts of Suggestion, Imitation, and Modeling in Reference Group Theory

In this theory, the individual is conceived of as influenced in behavior and attitude by his reference group (Schuman and Johnson 1976; Hyman and Singer 1968; Sherif and Sherif 1969; Merton and Rossi 1967; Merton 1967). As Jahoda has noted, reference-group theorists typically overlook a very central issue. They fail to "raise the question as to the conditions under which *different* reference groups become factors in determining opinions. As far as I know, this central question of reference group theory has not yet been tackled in research. As a rule, the concept is used as if the individual had only *one* group with whom he shared norms, namely the one deliberately built into the research design" (Jahoda 1972, my emphasis).

The concepts of suggestion, imitation, and modeling may provide a partial solution to the dilemma posed by Jahoda. One reference group may be chosen over another if a member of that reference group is publicized by the media. Thus, for example, the reference group "persons with troubles" may be chosen over others when the media publicize a story about a troubled person who commits suicide. On the other hand, the reference group "Christians" may become salient if the media carry a story about Billy Graham or the Pope, while still other reference groups may become salient on Father's Day, Mother's Day, George Washington's birthday, Jefferson Davis's birthday, or Martin Luther King's birthday.

Here again, as in our analysis of decision theory, publicity and suggestion may serve as mediating mechanisms helping to switch an individual from one choice to another—in this case from one reference group to another.

Possible Usefulness of the Concepts of Suggestion, Imitation, and Modeling in the Theory of Anomie and Its Effects

Many different behaviors have been thought to result from anomie. These range from various types of anomie-reducing social movements (Kornhauser 1959; Cantril 1963; Toch 1965; Klapp 1969; De Grazia 1948) to various types of deviant behavior, such as suicide, delinquency, alcoholism, drug addiction, and mental disorders (Durkheim 1951; Gibbs and Martin 1964; Merton 1967; Clinard 1964; Cloward 1959; Cloward and Ohlin 1960; Seeman 1975; for criticisms of this literature, see Clinard 1968; Gibbs and Erickson 1975; Akers 1973).

A central question in this literature has seldom been broached and has *never been resolved: Given that so many behaviors are thought to result* from anomie, what makes an individual respond to anomie with one type of behavior rather than another? Why do some individuals respond to anomie by committing suicide, while others turn to alcohol and still others join anomie-reducing social movements?[26]

The concepts of suggestion, imitation, and modeling may provide a partial answer to this question. An anomic individual may choose suicide rather than some other solution if a model individual has recently committed suicide—thus bringing suicide to the forefront as a salient solution to anomie.[27] On the other hand, if some alternative solution to suicide were to be made salient, then this alternative might be chosen, instead. Toch (1965, p. 124) provides an excellent example of this when he presents the case history of a woman "switched" from suicide to an alternative solution. "One day [the woman writes] I was sitting in my home and the thought came to me, 'Why don't you put your head in the oven, turn on the gas, and end it all?' Just

[26] Merton attempted a partial answer to this question when he noted that different types of structural conditions might lead to different types of deviant behavior. His approach is incomplete: it helps us to understand why an individual might be prone to retreatist rather than innovative behaviors, but it does not help us to understand why an individual might be prone to one type of retreatist behavior (say suicide) rather than another (say alcoholism). Similarly, his approach does not help us to understand why some individuals respond to anomie with deviance while others respond by joining anomie-reducing social movements. Subsequent modifications of Merton's paradigm (e.g., Cloward 1959; Cloward and Ohlin 1960) can also be shown to provide an incomplete answer to the question under discussion.

[27] Earlier I discussed the imitation of a general mode of problem solving, e.g., using the criteria of decision theory. Here I am concerned with a different type of imitation—the direct imitation of a particular, concrete solution, rather the general imitation of a method of reaching a solution.

then a knock came on the door. I answered, and saw a tall woman. . . . She asked if she could give me Bible studies. I said Yes . . . I was invited to attend some Adventist evangelist meetings. Later I was baptized. For me the Adventist faith is all that I need. It provides an inner peace that passes all understanding."

Here again, as in the analysis of decision theory and of reference group theory, suggestion may serve as a mediating mechanism helping to switch an individual from one choice to another—in this case from a type of deviant behavior to membership in an anomie-reducing religious movement.

I have now considered very briefly some ways one might apply the concepts of imitation, suggestion, and modeling to three areas of social theory. These three theories can be enriched and enlarged if they are extended to incorporate the processes of suggestion, imitation, and modeling. As we become sensitized to processes like suggestion we become aware of very general questions we might otherwise have missed. What switches an individual from one choice to another—from one problem to another, from one mode of solving problems to another, from one reference group to another, from one form of deviance to another? A partial answer to these questions may be found in the concepts of suggestion and imitation, which may serve as very general switching mechanisms in social behavior. If suggestion and imitation do indeed serve these very broad functions, we may find it useful to incorporate the concepts of suggestion and imitation into many theories of social behavior, not just the three theories discussed above.

REFERENCES

Akers, R. 1973. *Deviant Behavior: A Social Learning Approach.* Belmont, Calif.: Wadsworth.

Ayer Press. 1966–73. *Ayer Directory of Publications.* Philadelphia: Ayer Press.

Bandura, A. 1973. *Aggression: A Social Learning Analysis* Englewood Cliffs, N.J.: Prentice-Hall.

Berkowitz, L. 1962. *Aggression: A Social Psychological Analysis.* New York: McGraw-Hill.

Britt, S. H., ed. 1970. *Psychological Experiments in Consumer Behavior.* New York: Wiley.

Brown, P. L., and R. F. Birdie. 1960 "Driver Behavior and Scores on the MMPI." *Journal of Applied Psychology* 44:18–21.

California Highway Patrol. 1966–73. *Annual Report of Fatal and Injury Motor Vehicle Traffic Accidents.* Sacramento: California Highway Patrol.

Cantril, H. 1963. *The Psychology of Social Movements.* New York: Wiley.

Clinard, M., ed. 1964. *Anomie and Deviant Behavior.* New York· Free Press.

———. 1968. *Sociology of Deviant Behavior.* New York: Holt, Rinehart & Winston.

Cloward, R. A. 1959. "Illegitimate Means, Anomie, and Deviant Behavior." *American Sociological Review* 24:164–76.

Cloward, R. A., and L. E. Ohlin. 1960. *Delinquency and Opportunity.* New York: Free Press.

Cohen, J. B., ed. 1972. *Behavioral Science Foundations of Consumer Behavior.* New York: Free Press.

Comstock, G. 1975. *Television and Human Behavior: The Key Studies*. Santa Monica, Calif.: Rand.

Comstock, G., and M. Fisher. 1975. *Television and Human Behavior: A Guide to the Pertinent Scientific Literature*. Santa Monica, Calif.: Rand.

Comstock, G , and G. Lindsey. 1975. *Television and Human Behavior: The Research Horizon, Future and Present*. Santa Monica, Calif.: Rand.

Conger, J. J., H. S. Gaskill, D. D. Glad, R. V Rainey, W. L. Sawrey, and E. S. Turrell. 1957. "Personal and Interpersonal Factors in Motor Vehicle Accidents." *American Journal of Psychiatry* 113:1069–74.

———. 1959. "Psychological and Psychophysiological Factors in Motor Vehicle Accidents." *Journal of the American Medical Association* 169:1581–87.

Crancer, A., and D. L Quiring. 1970 "Driving Records of Persons Hospitalized for Suicidal Gestures " *Behavioral Research in Highway Safety* 1:33–42.

De Grazia, S. 1948. *The Political Community: A Study of Anomie*. Chicago: University of Chicago Press.

Dirksen, C. J., and A. Kroeger. 1973. *Advertising Principles and Problems*. Homewood, Ill.: Irwin.

Durkheim, É 1951. *Suicide*. Translated by J. A. Spaulding and G. Simpson. New York: Free Press.

Edwards, W., and A Tversky, eds. 1967. *Decision Making*. Baltimore: Penguin.

Engel, J. F., D. T. Kollat, and R. D. Blackwell. 1973. *Consumer Behavior* New York: Holt, Rinehart & Winston.

Finch, J. R., and J. P. Smith. 1970. *Psychiatric and Legal Aspects of Automobile Fatalities*. Springfield, Ill.: Thomas.

Flanders, J. P. 1968. "A Review of Research on Imitative Behavior." *Psychological Bulletin* 69·316–37.

Gibbs, J. P , and M L. Erickson. 1975. "Major Developments in the Sociological Study of Deviance." Pp. 21–42 in *Annual Review of Sociology*, edited by A. Inkeles, J. Coleman, and N Smelser. Palo Alto, Calif.: Annual Reviews.

Gibbs, J. P., and W T. Martin. 1964. *Status Integration and Suicide: A Sociological Study*. Eugene: University of Oregon Press.

Goranson, R. E. 1969. "A Review of Recent Literature on Psychological Effects of Media Portrayals of Violence." Pp. 395–413 in *Violence and the Media: A Staff Report to the National Commission on the Causes and Prevention of Violence*, edited by R. K. Baker and S J. Ball. Washington, D C.: Government Printing Office.

Hays, W. L. 1963. *Statistics*. New York: Holt, Rinehart & Winston.

Hoel, P. G. 1965. *Introduction to Mathematical Statistics*. New York: Wiley.

Hyman, H. H., and E. Singer, eds. 1968. *Readings in Reference Group Theory and Research*. New York: Free Press.

Jacoby, J. 1976. "Consumer Psychology: An Octennium." Pp. 331–58 in *Annual Review of Psychology*, edited by M. R. Rosenzweig and L. W. Porter. Palo Alto, Calif.: Annual Reviews.

Jahoda, M. 1972. "Conformity and Independence: A Psychological Analysis." Pp. 339–54 in *Behavioral Science Foundations of Consumer Behavior*, edited by J. B. Cohen. New York: Free Press.

Klapp, O. E. 1969. *Collective Search for Identity*. New York: Holt, Rinehart & Winston.

Klapper, J. T. 1960. *The Effects of Mass Communication*. New York: Free Press.

Kornhauser, W. 1959. *The Politics of Mass Society*. New York: Free Press.

Le Bon, G. 1895. *The Crowd*. London: T. Fisher Unwin.

Liebert, R. M., and N. S. Schwartzberg. 1977. "Effects of Mass Media." Pp. 141–74 in *Annual Review of Psychology*, edited by M. R. Rosenzweig and L. W. Porter. Palo Alto, Calif.: Annual Reviews.

Litman, R. E., and N. Tabachnick. 1967. "Fatal One-Car Accidents." *Psychoanalytic Quarterly* 36:248–59.

MacDonald, J. M. 1964. "Suicide and Homicide by Automobile." *American Journal of Psychiatry* 121:366–70.

Mead, G. H. 1934. *Mind, Self and Society: From the Standpoint of a Social Behaviorist.* Chicago: University of Chicago Press.

Merton, R. K. 1967. *Social Theory and Social Structure.* New York: Free Press.

Merton, R. K., and A. S. Rossi. 1967 "Contributions to the Theory of Reference Group Behavior." Pp. 225–80 in Merton 1967.

Michman, R. D., and D. W Jugenheimer, eds. 1976 *Strategic Advertising Decisions: Selected Readings.* Columbus, Ohio: Grid

Mosteller, F., R. E K Rourke, and G. B Thomas. 1970. *Probability with Statistical Applications.* Reading, Mass : Addison-Wesley.

Phillips, D. P. 1974. "The Influence of Suggestion on Suicide Substantive and Theoretical Implications of the Werther Effect " *American Sociological Review* 39·340–54.

———. 1977*a*. "Motor Vehicle Fatalities Increase Just after Publicized Suicide Stories " *Science* 196·1464–65.

———. 1977*b*. "Motor Vehicle Fatalities and the Mass Media " Paper presented at the Annual Meetings of the Population Association of America, St Louis, April 1977.

———. 1978 "Airplane Accident Fatalities Increase Just after Stories about Murder and Suicide " *Science* 201:748–50.

Phillips, D. P., and K A Feldman. 1973. "A Dip in Deaths before Ceremonial Occasions: Some New Relationships between Social Integration and Mortality " *American Sociological Review* 38·678–96.

Porterfield, A. L 1960. "Traffic Fatalities, Suicide and Homicide " *American Sociological Review* 25:897–901.

Preston, C. E. 1964 "Accident Proneness in Attempted Suicide and in Automobile Accident Victims." *Journal of Consulting Psychology* 28:79–82

Preston, C. E., and S Harris 1965. "Psychology of Drivers in Traffic Accidents." *Journal of Applied Psychology* 49·284–88

Raiffa, H. 1970. *Decision Analysis.* Reading, Mass : Addison-Wesley

Sandage, C. H , and V. Fryburger 1975 *Advertising Theory and Practice* Homewood, Ill : Irwin.

Schuman, H , and M. P. Johnson, eds. 1976 "Attitudes and Behavior." Pp 161–208 in *Annual Review of Sociology,* edited by A Inkeles, J. Coleman, and N. Smelser. Palo Alto, Calif.: Annual Reviews

Seeman, M. 1975. "Alienation Studies" Pp. 91–124 in *Annual Review of Sociology,* edited by A. Inkeles, J. Coleman, and N. Smelser Palo Alto, Calif : Annual Reviews

Selzer, M. L. 1969. "Alcoholism, Mental Illness, and Stress in 96 Drivers Causing Fatal Accidents " *Behavioral Science* 14·1–10.

Selzer, M. L., and C. E. Payne. 1962. "Automobile Accidents, Suicide, and Unconscious Motivation." *American Journal of Psychiatry* 119:237–40.

Sherif, M., and C. W. Sherif. 1969. *Social Psychology.* New York: Harper & Row.

Siegel, S. 1956. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Singer, J. L. 1971. "The Influence of Violence Portrayed in Television or Motion Pictures upon Overt Aggressive Behavior " Pp. 19–60 in *The Control of Aggression and Violence: Cognitive and Psychological Factors,* edited by J. L. Singer. New York: Academic Press.

Slovic, P., B Fischhoff, and S. Lichtenstein. 1977. "Behavioral Decision Theory." Pp. 1–40 in *Annual Review of Psychology,* edited by M. R. Rosenzweig and L. W. Porter. Palo Alto, Calif.: Annual Reviews.

Surgeon General's Scientific Advisory Committee on Television and Social Behavior. 1972. *Television and Growing Up: The Impact of Televised Violence.* Washington, D.C.: Government Printing Office.

Tabachnick, N., J. Gussen, R E. Litman, M L. Peck, N. Tiber, and C. I Wold. 1973. *Accident or Suicide?* Springfield, Ill.: Thomas.

Tabachnick, N., R. E. Litman, M. Osman, W. L. Jones, J. Cohn, A. Kasper, and J. Moffat. 1966. "Comparative Psychiatric Study of Accidental and Suicidal Death." *Archives of General Psychiatry* 14:60–68.

Tarde, G. 1903. *The Laws of Imitation.* New York: Holt.

Television Digest. 1970. *Television Fact Book*. Washington, D C.: Television Digest.
Tillman, W. A., and G. E. Hobbs. 1949. "The Accident-prone Automobile Driver: A Study of the Psychiatric and Social Background." *American Journal of Psychiatry* 106:321–31.
Toch, H. 1965. *The Social Psychology of Social Movements*. Indianapolis: Bobbs-Merrill.
U S. Public Health Service. 1974. *Vital Statistics of the U.S.* 1970. Vol. 2, pt. A. Washington, D.C.: Government Printing Office.
Williams, J. D. 1954. *The Compleat Strategyst*. New York: McGraw-Hill.

Socioeconomic Status and Mental Disorder: New Evidence and a Sociomedical Formulation[1]

William A. Rushing and Suzanne T. Ortega
Vanderbilt University

Despite numerous studies which show that socioeconomic status and mental illness are inversely related, it is possible that this relationship exists for some disorders but not others. Study of approximately 10,000 first admissions to state hospitals in one state covering the period 1956–65 shows an inverse relationship only for organic and schizophrenic disorders. In light of the etiological and predisposing factors that are involved in these two disorders, physical and medical factors appear to be more important than psychological (e.g., stress) factors in mediating the relationship with socioeconomic status. It is concluded, therefore, that a sociomedical formulation (in contrast to a sociopsychiatric formulation) is a plausible interpretation of the relationship between socioeconomic status and mental disorder.

Over the years a large number of studies of mental illness have shown that socioeconomic status and mental illness are negatively related (Fried 1969; Dohrenwend and Dohrenwend 1974). A question remains, however, as to whether there are differences by type of disorder. If differences do exist, they would indicate that a specification of the relationship is in order, and this would have implications for the interpretation of the causal role of socioeconomic factors in mental illness. Most interpretations emphasize the mediating role of sociopsychiatric factors, especially socioeconomically induced stress, stemming variously from status frustration, denied opportunity, and unfulfilled aspirations (see Dohrenwend and Dohrenwend 1974, p. 441). In contrast, we will present a causal interpretation in which sociomedical variables are emphasized.

Data are for all 21- to 64-year-old first admissions who were admitted and discharged during the 10-year period between fiscal 1956 and 1965 to each of the six Tennessee state mental hospitals. Diagnostic distinctions are made between brain disorders (due to toxins and organic impairments) and functional disorders, and within the latter, among schizophrenia, manic-depres-

1 We are grateful to George Becker, Jack P. Gibbs, Walter R Gove, David Mechanic, and especially to an anonymous referee who read and commented on this paper in its various stages of development. Comments by Raymond Duff, August B. Hollingshead, and Jerry Myers to an oral presentation were also helpful. Finally, we owe especial debts of gratitude to Jack Esco, who worked on constructing computer data files in the early stages of the research, and to Jack Holladay of the Tennessee Department of Mental Health and Mental Retardation, who made the data available.

sion, neuroses, and personality disorders. Of the total of 10,673 patients admitted for mental illness, 9,513 were classified in one of these categories.

FINDINGS

Educational Status

Hospitalization rates are obtained for each of the seven educational categories (0–4, 5–7, 8, 9–11, 12 years, some college, and college graduate and beyond) used to index socioeconomic status. Age-standardized ratios for all patients, separately for males and for females, are calculated for these categories, as follows[2]—males: 119, 110, 113, 103, 85, 77, and 48; females: 111, 108, 122, 107, 89, 61, and 55.

Figure 1 plots the ratios for brain disorders, schizophrenia, and all other disorders (referred to hereafter as "other"). The two highest categories (some college and college graduate and beyond) are combined (with the midpoint set at 14) because of small numbers. For both sexes the inverse relationship is specific to the brain and schizophrenic disorders, with a curvilinear pattern for the other category (manic-depression, neuroses, and personality disorders). Each of the three other disorders exhibits some deviation from the overall trend, although each disorder manifests the same general pattern (data not shown).

Occupational Status

On entering a hospital, patients were classified in one of the occupational categories in the Dictionary of Occupational Titles (DOT) of the Department of Labor. Many of these categories are identical to the detailed occupational categories (DOC) of the U.S. census, and several others are virtually so. Socioeconomic index (SEI) scores are available for the DOC titles (Duncan 1961). The analysis is limited to males.

Since the number of patients (and the population) for some occupations is quite small, a change in just one or two patients could lead to a substantial change in the hospitalization rate. To correct for this (and to follow the

[2] The method for calculating the ratio is the same as that used by the Department of Health, Education, and Welfare in computing mortality ratios (Guralnick 1963*a*, 1963*b*) and is the ratio of tabulated cases in an educational category to the expected number if the rate for each of the several educational categories were the same as for the total sample. The ratio is then multiplied by 100 (The ratio is not to be confused with the rate, which would indicate the number of persons in each educational category who were hospitalized per unit of population) Rates are based on 8,568 patients from 25 to 64 years old since the census for states does not provide population figures for the 21–24 year olds by education (U.S. Bureau of the Census 1963, pp. 315–16). However, crude rates including 21–24 year olds give the same relationships as those for the age-standardized ratios.

TABLE 1

SOCIOECONOMIC SCORE, HOSPITALIZATION RATES PER 100,000 MALES, AND NUMBER OF MALE ADMISSIONS FOR 23 OCCUPATIONAL CATEGORIES—TENNESSEE, 1956–65

OCCUPATION*	SOCIO-ECONOMIC STATUS SCORE	HOSPITALIZATION RATES PER 100,000			NO OF ADMISSIONS
		Brain	Schizo-phrenia	Other	
Teachers (primary, secondary, and n.e.c.)†	78	26	207	194	34
Insurance salesmen‡	76	75	15	313	27
Clergymen	71	205	205	225	32
Electricians	69	246	77	447	50
Stock clerks§	69	116	70	381	21
Salespersons and clerks‖	67	68	92	329	123
Plumbers, gas fitters, and steam fitters#	66	231	71	266	33
Machinists	65	103	88	250	30
Brick and stone masons and tile setters	62	106	42	424	27
Mechanics and repairmen, n e.c.	62	52	52	165	57
Welders and flame cutters	61	67	134	370	35
Cranemen, derrickmen, hoistmen, and shovelmen**	59	327	109	850	23
Carpenters	58	200	146	305	144
Mechanics and repairmen (motor vehicles)	58	193	171	520	120
Attendants (filling station and parking lots)††	58	144	101	278	48
Painters (construction and maintenance)	56	413	224	508	97
Cooks (except private family)	54	206	103	378	20
Farmers	53	307	216	333	731
Attendants (hospitals and other institutions, n.e.c.)	52	206	124	578	22
Mining and mine machine operators	49	294	52	225	33
Janitors and sextons	47	101	71	131	30
Farm workers	42	129	287	90	143
Porters‡‡	36	159	176	106	30

SOURCE —For population base (14 years and above) for calculating hospitalization rates, U S. Bureau of the Census (1963· 411–16)

* Unless otherwise indicated, occupational title is as it appears in both the DOT and DOC classification systems.

† Includes from DOT, "teachers, primary school and kindergarten," "teachers (secondary school) and principals," and "teachers and instructors, n.e c.", from DOC, "teachers elementary schools," "teachers secondary schools," and "teachers, n e c."

‡ Listed as "insurance agents, brokers, and underwriters" in DOC

§ Listed as "stock clerks and storekeepers" in DOC

‖ Includes from DOT, "sales clerks," "sales persons," "salesmen, to consumers," and "sales clerks, dry cleaning and laundry"; from DOC, "salesmen and sales clerks, retail trade "

\# Listed as "plumbers and pipe fitters" in DOC

** Listed as "cranemen, derrickmen, and hoistmen" in DOC

†† Listed as "attendants, auto service and parking" in DOC

‡‡ Includes from DOT, "pullman porters," "baggage porters," and "porters, n e c."; from DOC, "porters "

procedure used below in the section entitled Evidence for Occupational and Industrial Accidents), all occupations with fewer than 20 admissions are excluded. All residual categories for major occupational and industrial categories (e.g., "professional occupations, not elsewhere classified [n.e.c.]") are also eliminated. This leaves 23 DOT categories which are identical or virtually identical to DOC categories (see table 1).[3] Although the occupations are

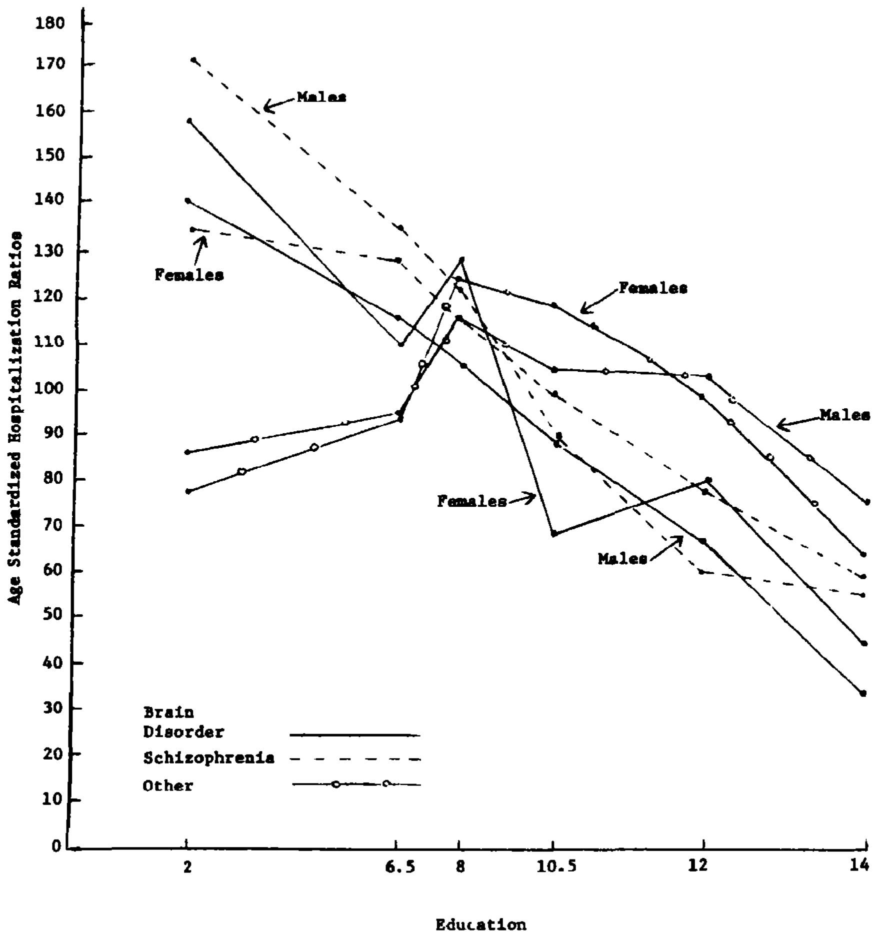


FIG. 1.—Plot of age-standardized hospitalization ratios on midpoint of educational categories—Tennessee, 1956–65.

[3] In seven cases, however, the DOT category has two or three levels (skilled, semiskilled, and unskilled), whereas the DOC category has only one. One category—cranemen, derrickmen, hoistmen, and shovelmen—has three levels, and six categories—auto mechanics and repairmen; mechanics and repairmen (not elsewhere classified); electricians; mining and machine operators; painters, construction and maintenance; and welders and flame cutters—have two levels (skilled and semiskilled). Since most of

clearly not a systematic or random sample (most are manual and service occupations), socioeconomic values from all but the highest levels are represented; and all levels are represented from the blue collar and service categories from 69 for electricians to 36 for porters.

The Spearman rank correlation (ρ) is substantially higher for brain disorders (—.37) and schizophrenia (—.42) than for other disorders (.14). Thus results for occupational status also indicate that the negative relationship of socioeconomic status to mental illness is specific to brain disorders and schizophrenia.[4] Consequently, interpretation of the relationship between

---

these occupations require considerable skill, the rule of procedure was to include only those admissions from the highest skill category listed for the occupation in the DOT system. For example, only admissions from the cranemen, et al. category classified as skilled were included in this category, thus excluding all semiskilled and unskilled cranemen, et al.; for welders and flame cutters, only two categories are listed—semiskilled and unskilled, so all the unskilled welders and flame cutters are eliminated. In all cases, however, the rule led to the elimination of very few cases. (Two of 122 for auto mechanics and repairmen; two of 98 for painters, construction and maintenance; one of 51 for electricians; three of 38 for welders and flame cutters; three of 35 for mining and mine machine operators; six of 61 for mechanics and repairmen (not elsewhere classified); and four of 27 for cranemen, derrickmen, hoistmen, and shovelmen.) We believe that the elimination of the semi- and unskilled admissions for occupations with three skill levels is warranted because all of these occupations require considerable skill, even those which are classified in the DOC as in the operative rather than craftsmen category.

[4] Our results are consistent with nearly all the evidence. In his review of the studies of socioeconomic status and schizophrenia, Kohn (1968) reports that almost every study reports a negative relationship (see also Kohn 1976). In their reviews of "true" prevalence studies, the Dohrenwends (1969, 1974) also report that most studies show a negative relationship for schizophrenia but no particular relationship for neurosis and manic-depression. Although sociologists and psychiatrists have generally neglected brain disorders, the limited evidence indicates a negative relationship with class status (Hollingshead and Redlich 1958, p. 233; Jaco 1960, pp. 210–27; see also Langner and Michael 1963, pp. 403, 407) On the surface, findings for personality disorder in this study appear to differ from most studies. The Dohrenwends (1969, p 30) report that in nine of 12 studies, the maximum rate is reported for the lowest socioeconomic stratum. Comparison of our results and these studies is problematic, however. First, the measurement of social class is imprecise in several of the studies; second, all are studies of prevalence rather than of incidence; and third, all were published between 1931 and 1964. Date of publication is significant since the diagnostic classification system used for patients in this study was developed in 1952 by the American Psychiatric Association (APA) (Committee on Nomenclature and Statistics of the American Psychiatric Association 1952). This classification includes five major categories under personality disorders and 19 specific subcategories which include emotionally unstable personality, passive-aggressive personality, sexual deviation, and alcoholism Prior to 1952, the diagnosis of personality disorder was often limited to "psychopathic personality" (this is replaced with "sociopathic personality" in the APA classification). In at least four studies cited by the Dohrenwends, this is indeed the case (Hyde and Kingsley 1944; Bremer 1951; Lin 1955; Helgason 1964) In another study, personality disorders are limited to "personality trait" disorders (one of the five major categories for personality disorders in the APA classification scheme), which Langner and Michael admit "lacked clear definition" in the research and included disorders covering a "wide range of . . . mental mechanisms" and etiologies (1963, p. 56). Furthermore, although they do report

socioeconomic status and mental illness must include only those class-related factors involved in the production of these two disorders.

BRAIN DISORDERS

Much of the debate among sociologists, psychiatrists, and others concerns the applicability of the medical model to mental disorders. Sociologists (e.g., Goffman 1961; Scheff 1966, 1974; and Gove 1976, p. 565) contend that it is not valid when applied to mental illness, and some psychiatrists hold the same view (e.g., Szasz 1961; Torrey 1974). The general argument is "that those disorders we call 'mental illnesses' are different in crucial respects from physical illness: they do not have the same biological existence and therefore cannot be understood within the causal framework of natural science [as physical illnesses are understood]. Whereas physical illness is seen as an objective deviation from the norms of physical structure and function, 'mental illnesses' are more properly described as infractions of social norms" (Morgan 1975, p. 268). Such a conclusion, however, would not apply to brain disorders, in which case the medical model is valid (Morgan 1975, p. 265; Torrey 1974, p. 155).

Organic trauma or pathogens are, of course, the precipitating, proximal, or immediate causes of brain disorders. And the pathogens are diverse depending on the specific disease. Consequently, an explanation for the relationship between socioeconomic status and brain disorders must include disorders which have *diverse* physical causes. We suggest there are three general sets of variables: the physical environment, health attitudes and behavior, and medical care.

---

a higher rate of these disorders in the lowest socioeconomic stratum (p. 402), there is no tendency for specific symptoms associated with sociopathic personality disorders to be more frequent in the lowest stratum: alcoholic and dyssocial symptomatology were greater in the lowest stratum, but sexual deviance and aggresiveness were three to five times more frequent in the higher strata (p. 407). In still another study (Gillis, Lewis, and Slabbert 1968) figures are given for "symptom patterns" rather than cases, a point the Dohrenwends note (1969, p. 30). Finally, in three studies the differences between the rate for the lowest stratum and the stratum having the minimum rate is small (percentages of .59, .27, and .12) (see Dohrenwend and Dohrenwend 1969, p. 30). Consequently, a change in the socioeconomic classification of very few patients could yield different results. For example, if just one patient were classified from "lower class" to "upper class" in one study, the "upper class" would have the maximum rate (Lin 1953, p. 327). Hence, with very small changes (all of which could easily occur because of measurement or sampling errors), 50% of the studies might report the maximum rate for some stratum other than the lowest. It would seem, therefore, that an inverse relationship between socioeconomic status and personality disorder has not been firmly established. In any case, with the possible exception of personality disorder, findings for the present study are entirely consistent with well-established patterns in research on socioeconomic status and mental disorder. They are also consistent with evidence for most differences by sex (see the review by Gove and Tudor 1973; also Dohrenwend and Dohrenwend 1977) and marital status (Gove 1972). Differences by sex and marital status are topics of forthcoming papers.

Physical Environment

We include here hazards that are apt to contribute to accidents (e.g., occupational hazards) as well as the more general cultural and social components of the environment (e.g., human aggression). Evidence shows that accidents and injuries to adults are more frequent in the lower socioeconomic strata (Sanders 1964, p. 670). Also, low-status occupations are disproportionately manual and hence more physically hazardous than high-status jobs. Although some authors question whether aggressive behavior is more frequent in the lower socioeconomic strata (see Straus 1971; Steinmetz 1974), others believe that aggression is negatively related to socioeconomic status (see Miller 1958; Goode 1971; and Banfield 1974, pp. 61–64). Most evidence seems to support the latter position. Lower-status persons (classes 4 and 5 on Hollingshead's Two-Factor Index) are more apt to file complaints of physical abuse in divorce proceedings (Levinger 1966); lower status parents are more apt to use physical punishment (see Gold 1958; Bronfenbrenner 1958; and Kohn 1969); and children in lower socioeconomic, poverty, and welfare families are more apt to be victims of child abuse (see Gil 1970; and Erlanger 1974). In addition, anecdotal and ethnographic accounts would indicate support for this belief (see Winslow 1968, pp. 122–23).[5]

Although physical accidents and violence only rarely result in brain damage, an increase in their frequency increases the probability of damage to the body including the brain. Consequently, the fact that physical accidents and violence are more frequent in the lower socioeconomic strata increases the probability of injuries to the brain.[6]

Health Attitudes and Behavior

A review of studies of health-relevant attitudes and behavior indicates that those typical of both lower-strata persons and health-care providers cause

[5] The following comments from teenage slum dwellers illustrate this (Winslow 1968, pp. 122–23). "When I first started living around here it was really bad, but I have gotten used to it now. Been here two years. People getting beat up . . Gee, there's a lot of violence around here. You see it all the time. . . ." "Sometimes where I live at people be hitting at each other, fighting next door Then when they stop fighting then you can get some sleep. . . ." "Drinking, cussing, stabbing people, have policemen running all around mostly every day in the summertime." (See also Toby 1966.) Two studies which raise questions concerning the negative relationship between socioeconomic status and violence are attitude and value studies of college students (Straus 1971; Steinmetz 1974); as a result the socioeconomic scale is necessarily truncated at the lower end.

[6] Note that this formulation stipulates that mental illness is an outgrowth of violence rather than the reverse. Others who observe that violence and mental illness are more frequent in the lower socioeconomic strata view mental illness as a cause of violence. Banfield, for example, states that while "lower class life [is] extraordinarily violent, . . . much of the violence is probably more an expression of mental illness than of class culture" (Banfield 1974, p. 63). Our position is that the higher rate of mental illness in the lower class is due, in part, to a higher rate of violence.

lower-status persons to use health services less than is medically indicated (Berkanovic and Reeder 1974). Such persons know less about the treatability and severity of symptoms, have less confidence in the efficacy of medical treatment, and make less use of preventive and other discretionary services (Ross 1962; Richardson 1972; Rosenstock 1969; Bice, Eichhorn, and Fox 1972). Poor personal hygiene and failure to take necessary preventive measures to avoid contracting a disease (e.g., syphilis) may be related to class status. Also, excessive drinking is probably not as severely sanctioned in the lower strata and the dangers of prolonged drinking may not be as well known. Thus, attitudes and behavior as aspects of cultural and social life-style may be contributing factors to the higher rates of brain disorder in the lower socioeconomic strata.

Medical Care

The bulk of the evidence on physical illness indicates that the rates of mortality and morbidity are higher in the lower socioeconomic strata (Antonovsky 1972; Mechanic 1968, pp. 236–61; Richardson 1972). Since members of the lower strata apparently do not now receive a disproportionately lower proportion of health care than higher socioeconomic groups (see reviews by Bice et al. 1972; Anderson and Anderson 1977), the higher illness rates in the lower strata may be due to other factors such as public sanitation, personal hygiene, and health practices and attitudes. Some evidence suggests, however, that once need for medical care is controlled, physician contacts may still be positively related to socioeconomic status (Miles 1977). More significantly, even though socioeconomic differences in medical care may be small today, substantial differences existed before the patients in this study were hospitalized (1956–65) (Bice et al. 1972; Anderson and Anderson 1977). Moreover, lower-status persons are more apt to receive care in hospital clinics and emergency rooms and to be set apart when they are hospitalized; they are less apt to have regular physicians and more apt to be seen by chiropractors, foreign-educated physicians, or physicians' assistants (Department of Health, Education, and Welfare 1972; Herman 1972; Roth 1969; Duff and Hollingshead 1968; Alpert et al. 1969; Studnicki, Saywell, and Wiecheter 1976; Koos 1954; Rushing and Miles 1977). These differences in types of care, and differences in preventive care may contribute to failures to detect certain conditions that may subsequently result in brain damage: epilepsy, Huntington's chorea, and pernicious anemia, all of which may be associated with schizophrenia (see Davidson and Bagley 1969; and Brune 1971) as well as syphilis, brain tumors, intracranial lesions, and possibly alcoholism. Also, the lack of a regular physician and hence continuity of care may contribute to intrauterine infection, congenital defects, birth

trauma, and other conditions that result in brain disorders; and evidence does show that both infant mortality and postneonatal mortality rates are higher in lower-status groups (Lerner 1969, pp. 94–100).

In summary, the inverse relationship between socioeconomic status and brain disorders may be due in large part to such things as the physical environment, health attitudes and behavior, and access to quality medical care. Evidence for occupations and industry is consistent with the formulation as it pertains to the physical environment.

EVIDENCE FOR OCCUPATIONAL AND INDUSTRIAL ACCIDENTS

If socioeconomic factors which contribute to bodily injury in general contribute also to brain injury, occupations with higher rates of accidental injuries should also have higher rates of brain disorders in comparison with rates of functional disorders. Unfortunately, data are not available for accident rates by occupation. However, Guralnick (1963*a*, 1963*b*) reports 1950 male mortality rates for occupational categories included in the U.S. Census Intermediate Occupational Classification (IOC). The age-standardized mortality ratio (SMR) is reported for each category if there were at least 20 deaths. There are 13 categories that are the same or virtually the same as those in the list of 23 from the DOT system in table 1. These occupations (and 6 others) are presented in table 2.

For the 13 occupations, ρ between SMR and brain rate is .74 in comparison with .05 and .29 for the schizophrenic and other rates (for all functional disorders combined, $\rho = .29$). Since results could be affected by the small number of occupations, all occupations are included with reported SMRs if they had at least 10 hospital admissions, giving six additional cases. (As it turns out, all of these occupations have 15 or more admissions.) The ρ for the brain rate is .44 in comparison with .05 and .01 for the schizophrenic and other rates (.09 for all functional disorders). Results thus support the view that factors associated with injury to the body are contributing factors to impairment of the brain.

The DOT classifies a number of blue-collar manual occupations according to industries, seven of which are identical or approximately so to industries for which SMRs are reported by Guralnick (1963*b*) (see table 3). Schizophrenics and patients in the other category are combined because of the small number of cases for certain industry categories. Also, since state population is not given in the census for all categories in table 3, the ratio of brain to functional disorders is given rather than rates for the industrial population. Although there are few cases, only two of which permit exact comparisons, and most white-collar occupations are excluded, the correspondence in the ranking for the SMR and the ratio is clear; $\rho = .57$.

Results may be distorted because individual industries contain occupations that are heterogeneous in respect to physical danger and job-related accidental deaths. In one instance, however, data are available for the same occupation for different industries. Both the IOC and DOT classifications distinguish between "painters, construction and maintenance" and "painters, except construction." Since the SMR for the construction industry is relatively high (181), we would expect a higher brain to functional disorder ratio for painters in this industry than for other painters. Results show this to be the case. The ratio for nonconstruction painters is 0.167 (1:6) in comparison with 0.508 (32:63) for painters in the construction industry. Comparisons are based on painters classified in the skilled category in the DOT

TABLE 2

AGE-STANDARDIZED JOB-RELATED ACCIDENTAL MORTALITY RATIOS FOR ALL U.S. MALES, 20–64, 1950, AND HOSPITALIZATION RATES FOR TENNESSEE, 1956–65

Occupation*	Job-related Accidental Mortality Ratio (SMR)	Hospitalization Rates per 100,000			
		Brain	Schizophrenia	Other	No
Lumbermen, raftsmen, and woodchoppers	844	151	176	187	18
Mine operatives and laborers (n.e.c.)	581	294	225	52	36
Linemen and servicemen (telegraph, telephone, and power)	429	67	175	88	15
Cranemen, hoistmen, and construction machinery operators†	273	329	850	109	30
Electricians	231	246	447	77	52
Plumbers and pipe fitters‡	136	248	266	71	34
Painters (construction and maintenance)	131	413	508	224	102
Policemen, sheriffs, and marshalls§	130	168	223	28	17
Farmers and farm laborers	114	254	274	233	999
Welders and flame-cutters	107	67	370	134	40
Guards and watchmen‖	106	175	153	102	19
Carpenters	93	205	305	146	148
Civil engineers	88	111	295	148	15
Tinsmiths, coppersmiths, and sheet metal workers	85	202	404	202	16
Other mechanics and repairmen and loom fixers	77	56	165	51	65
Masons, tile setters, and stone cutters#	69	106	424	42	29
Mechanics and repairmen (automobile)	64	200	520	171	129
Janitors and porters**	49	101	131	71	61
Machinists	46	102	250	88	31

SOURCES.—Guralnick (1963b) for SMRs, and U.S. Bureau of the Census (1963) for population base (14 years and above) in calculating hospitalization rates.

* Occupational title is as listed in the IOC.

† Listed as "cranemen, derrickmen, hoistmen, and shovelmen" in DOT.

‡ Listed as "plumbers, gas fitters, and steam fitters" in DOT.

§ Includes "policemen and detectives (public and private)" and "sheriffs and bailiffs" from DOT

‖ Listed as "guards and watchmen, except crossing watchmen" in DOT.

\# Listed as "brickmasons, stonemasons, and tile setters" in DOT.

** Includes "janitors and sextons," "pullman porters," "baggage porters," "porters, n.e.c." from DOT.

TABLE 3

STANDARDIZED MORTALITY RATIO FOR JOB-RELATED ACCIDENTAL MORTALITY FOR 20- TO 64-YEAR-OLD MALES, UNITED STATES, 1950, AND RATIO OF BRAIN TO FUNCTIONAL DISORDER ADMISSIONS TO STATE MENTAL HOSPITALS IN TENNESSEE, 1956–65, BY INDUSTRY

Industry*	Age Standardized Mortality Ratio (SMR)	Brain to Functional Disorder Ratio	No of Hospital Admissions
Logging†	1,031	50	21
Mining‡	435	.81	38
Transportation, communication, and other public utilities	190	34	473
Construction	181	63	353
Stone and clay products§	133	55	14
Paper and allied products‖	79	00	6
Printing, publishing, and allied industries#	28	22	22

SOURCE.—For SMRs, Guralnick (1963*b*).
* Industry is listed as in Guralnick (1963*b*)
† Occupations included from the DOT classification are "inspectors, scalers, and graders, log and lumber" and "lumbermen, raftsmen, and woodchoppers"
‡ Occupations from the DOT classification are those listed under "occupations in extraction of minerals"
§ Occupations from the DOT classification are those listed under "occupations in production of stone, clay, and glass products."
‖ Occupations from the DOT classification are those listed under "occupations in production of paper and paper goods"
\# Occupations from the DOT classification are those listed under "printing occupations"

system. If semiskilled and unskilled are included (see n. 2), the comparisons are 0.09 (1:11) and 0.515 (33:64).[7]

Thus, as stated in our conception of physical environment, socioeconomic factors which contribute to physical illness and injury in general appear also to be involved in the etiology of toxic-organic mental disorders.

SCHIZOPHRENIA

Several biological and physiological factors are involved in schizophrenia that may be related to socioeconomic status. They are classified as genetics, biochemical and neurological factors, and viruses.

Genetics

A number of studies, twin studies in particular, provide compelling evidence that heredity is an important factor in schizophrenia (see Shields 1968;

[7] The category "painters, except construction and maintenance" is not included in table 1 because it did not have at least 20 admissions in the skilled level (Inclusion of all skill levels is appropriate when comparisons are among industries; their inclusion is questionable when comparisons are for occupations of different socioeconomic status.) Since the socioeconomic scores of the two painter categories are virtually the same (56 and 57), the industry comparisons would indicate that physical hazards associated with work are more crucial in the brain disorders than socioeconomic status per se.

Gottesman and Shields 1972; Campion and Tucker 1973; Kety 1974; Boklage 1977). As an explanation, some believe brain lateralization is important, with "the primary focus of schizophrenia brain dysfunction [being] found in the left hemisphere" (Boklage 1977, p. 32). Others observe that the hereditary factors which contribute to certain organic conditions may also be involved in schizophrenia (Taylor 1977; and Trimbel 1977). Significantly, behavior associated with hereditary schizophrenia is similar to that found in organic schizophrenia and in brain disorders with "schizophrenia-like" symptoms, such as epilepsy, especially temporal lobe epilepsy (Slater, Beard, and Glithero 1963; Taylor 1977, p. 404). But studies have yet to reveal precisely what the genetic factors are (e.g., see Shields 1968, pp. 111–17).

It is even less clear how genetic factors might account for the role of socioeconomic status in schizophrenia. One possibility, of course, is that (hereditary) schizophrenics have greater difficulty performing high-status occupational roles. Consequently, they "drift" downward (or remain at their beginning points) in the class structure. On balance, however, a review of the evidence would indicate that drift is not the sole explanation (Kohn 1968, pp. 158–60). We will have more to say about this hypothesis later.

For the present we will simply note that class differences in medical care provide an explanation that is as plausible as the drift hypothesis. Note first that it is not schizophrenia itself that is inherited. "Schizophrenic psychosis is clearly not what is inherited; that takes on the average near 30 years to appear, and it is not possible in the majority of cases to say that the individual 'was always more or less like that.' Therefore, what is inherited or otherwise cellularly imposed before birth is some cellular *anlage,* stable to 20–60 years of growth with or without intervening development of the *anlage* itself or of identifiable behavioral deviations originating therefrom" (Boklage 1977, p. 30). Taylor believes that "schizophrenic symptomatology results from disorganization of the structures subserving language and thought (and hence also socialization), normally compactly located in the left brain. *Developmental* anomalies which lead to alternative strategies in locating these skills lead to diffusion of the necessary association networks. In the *course of development* the neural base of language contracts first from hemispheric equipotentiality to the left hemisphere, and *later* contracts within the hemisphere to very localized areas" (1977, p. 406; emphasis added). In this view the onset of hereditary schizophrenia would vary with age and is precipitated by *developmental* anomaly: "Many different types of developmental anomaly may therefore lead to the deployment of a thinking and behavioral strategy which we call schizophrenia" (ibid.). Schizophrenia may also depend on damage incurred during brain development (ibid.).

If schizophrenia is not itself inherited but develops out of inherited pre-

dispositions, many cases might be prevented if proper medical care were provided. Indeed, Taylor notes that alert pediatricians may be able to identify symptoms which are precursors of schizophrenia later in life (1977, p. 406). Further, the risk of schizophrenia is disproportionately high for persons with temporal lobe epilepsy (Taylor 1975), and both schizophrenic symptoms and epileptic seizures are known to be influenced by the same drug, dopamine (Trimble 1977).

All this suggests that hereditary vulnerabilities or predispositions toward schizophrenia may be prevented by the intervention of good specialized medical care, which may be inversely related to class status. Therefore, the fact that genetic factors may be paramount in schizophrenia does not necessarily make environmental factors, such as socioeconomic status, any less important (for comments on the environment in general, see Shields 1968).[8]

Biochemical and Neurological Factors

Symthies (1975) and Trimble (1977) theorize that existing evidence indicates that biochemical factors are involved in schizophrenia. The evidence for neurological factors is substantial, with a number of conditions (e.g., nonlocalized neurological damage, irregular motor responses and muscle tone, EEG abnormalities, and central nervous system dysfunctions) found to be associated with schizophrenia (for a review see Campion and Tucker 1973, p. 461). Taylor considers the evidence "quite formidable" and so compelling that he wonders about the validity of the whole conception of functional psychoses (1977, pp. 405–6). It is not clear, however, whether the neurological deficits are due to heredity or to prenatal factors (Campion and Tucker 1973, p. 461). It is also not clear how they would be related to socioeconomic status. For example, schizophrenics and nonschizophrenics differ in the functioning of the extensor digitorum brevis muscle (Goode et al. 1977), but there is no reason to believe that the functioning of this

[8] Research indicates that genetic factors are also important in manic-depression (see Allen 1976 and Gershon, Dunner, and Goodwin 1971). We have seen, however, that manic-depression is not significantly related to educational status Although this appears to be consistent with reviews of research on socioeconomic status and manic-depression (see n. 4 above), the lack of a relationship may be due to the relatively small number of cases of manic-depressives in the present study (170 males and 308 females) Still another possibility is that different indices of socioeconomic status produce different results depending on the disorder; this issue is explored in a forthcoming paper. It is also conceivable that the genetic component in manic-depression does not interact with medical care, or at least with the class determinants of access to medical care, as we have suggested for schizophrenia. Significantly, in their review of the evidence concerning a genetic component in the affective psychoses, Gershon et al. (1971) do not even mention the possibility of interaction between socioeconomic factors and genetic components. However they do discuss the interaction of parental loss and parental deprivation with genetic factors as possibly significant in the etiology of depressive psychoses (1971, esp. pp. 5–6).

muscle varies by socioeconomic status. In other instances, however, linkages are indicated.

Evidence shows that individuals with brain disorders may exhibit symptoms similar to those of schizophrenia (Brackbill and Fine 1956; Davidson and Bagley 1969; Torrey and Peterson 1976). These disorders include Huntington's chorea, general paresis, pernicious anemia, and disorders with an epileptic base (for reviews see Davidson and Bagley 1969; Brune 1971). Evidence also shows that patients with brain tumors and persons who experience cerebral trauma frequently manifest schizophrenia-like symptoms, as do persons with lupus (Remington and Rubert 1962; Slater et al. 1963; Davidson and Bagley 1969; Klerman 1975, p. 86; Taylor 1975), and, as noted previously, epilepsy and schizophrenia appear to be related to a common biochemical mechanism since each responds to dopamine (Trimble 1977). Many of these factors could contribute to the inverse relationship between socioeconomic status and schizophrenia through two routes.

First, the effects of most, if not all, of the neurological impairments might be minimized with appropriate medical treatment, and access to high quality medical care and utilization of preventive care are positively related to socioeconomic status (see above). Even those who question whether the genetic interpretations are not overstated indicate the possible role of socioeconomic factors in neurological deficits. Specifically, Campion and Tucker (1973) believe that some of the results of twin studies could be accounted for by such factors as intrauterine insufficiency, perinatal hypoxia, and birth trauma. These conditions contribute to prenatal damage and infant mortality. The latter is known to be inversely related to socioeconomic status (Lerner 1969) and the former probably is. Consequently, Campion and Tucker (1973, p. 463) properly caution that information is needed about the class background of subjects in the twin studies. In some instances, of course, it may be questioned whether the schizophrenia-like thought and behavior really constitute schizophrenia (for a discussion, see Davidson and Bagley 1969, pp. 151–53). This would clearly be the case when behavior due to brain tumors and general paresis is diagnosed as due to schizophrenia. In these instances it would probably be correct to speak of the *mis*diagnosis of schizophrenia.[9] And this is the second socioeconomic route through which neurological impairments may operate. We have already seen that brain disorders are inversely related to socioeconomic status. Although the reliability of brain diagnoses would appear to be relatively high (see Torrey 1974, p. 67), errors do occur. Consequently, erroneous schizophrenic diag-

[9] The issue is further confused because schizophrenia is probably not a unitary disease with a common etiology but "is comprised of various disorders" with "multiple determinants in which there is not certainty as to the exact etiology" (Klerman 1975, p. 84). Nevertheless, some evidence would suggest that agreement among psychiatrists as to which patients are schizophrenic is relatively high (Kety 1974).

noses *typically* associated with brain disorders would also be related to socioeconomic status.

Viruses

A review of clinical, epidemiological, and laboratory evidence indicates that viruses are involved in the etiology of schizophrenia (Torrey and Peterson 1976). Although an understanding of this possible involvement is particularly difficult because of immunizations and the existence of "slow viruses" (viruses that remain latent and asymptomatic for periods of five to 20 years), evidence suggests that viral encephalitides, in particular, may be involved in schizophrenia (ibid.), and this has been noted for a number of years (Hendrick 1928, p. 1009). In addition, studies indicate unusually higher incidence of abnormal dermatoglyphics (fingerprints, palm prints, and footprints) and abnormal capillary formations in schizophrenics (Torrey and Peterson 1976, p. 139). Both conditions may result from intrauterine viral infections. It has also been reported that several studies in Sweden, Norway, Denmark, Great Britain, and the United States show a disproportionate number of schizophrenics being born in January through April, a period in which many viruses (e.g., rubella, measles, varicellazonter) show a similar peak of incidence (ibid., p. 140).

The vital hypothesis of schizophrenia is consistent with the findings for socioeconomic status on two counts. First, individuals can be immunized against many viral diseases and otherwise protected through preventive measures. While mass immunizations have certainly reduced socioeconomic differences, individuals in the lower strata are probably less apt to be immunized especially if it depends on purposive action since lower-class persons are less apt to be knowledgeable about the need for immunization. They are also less apt to recognize the value of preventive practices in general or to have a continuous relationship with a personal physician who would urge that preventive steps were taken. And lower-class mothers may receive lower quality prenatal care, thus increasing risks of intrauterine viral infections which affect the fetus.

Second, people must live in close and regular contact before viruses can be maintained in the community (Torrey and Peterson 1976, p. 140). Some viral diseases became prevalent only with the development of urbanization. Thus they may be most prevalent in the lower strata because of crowded living conditions, particularly in urban areas.

Summary

Socioeconomic factors may perform significant etiologic roles in brain disorders and schizophrenia in two ways. First, lower-status individuals may

be more apt to be exposed to etiological factors (e.g., infectious diseases) of both disorders. Second, socioeconomic factors may mediate the effect of biological factors (e.g., infectious diseases, neurological deficits). Many of the factors associated with brain and schizophrenic disorders may be modified by appropriate health-relevant attitudes and behavior (e.g., preventive care) and medical attention, which may vary by socioeconomic status. Therefore, rather than the relationship of a sociological variable with mental illness calling for the rejection of the medical model as is commonly supposed by sociologists, the integration of medical and sociological factors in a common (sociomedical) interpretation is called for instead.

ALTERNATIVE FORMULATIONS

Other interpretations of the relationship between socioeconomic status and mental disorder may be classified into three categories: drift or social selection, socioeconomic stress, and labeling.

Drift

For both brain disorders and schizophrenia, drift is controlled to some extent with the measure of education; psychiatric status probably has less effect on education than on occupation or income since the education of most persons is completed prior to adulthood. Still it might be argued that in some cases of brain disorder (e.g., such as congenital disorders and disorders due to birth trauma), the condition causes low educational achievement. This might be at least partially responsible for the high rate of brain disorders for patients with little or no education. However, there are only 26 cases with congenital and birth trauma diagnoses in the sample; their removal has no effect on the relationship in figure 1. Also, it is not clear why persons with childhood and adolescence disorders would not be hospitalized until after they were over 20 years of age. It is conceivable, of course, that many individuals are cared for by parents until adulthood, at which point they are placed in institutions. It this were the explanation for our results, unusually high rates would be expected for young adults with low educations. Table 4 shows, however, that this is not the case. With only a few exceptions, brain disorders and schizophrenia tend to decrease as education increases for all age categories.

In addition, the nature of the relationship would appear not to be consistent with a drift formulation. Figure 1 shows that for males, brain disorder and schizophrenic rates continuously decrease with educational status and, with only two exceptions for brain disorders, the same is true for females. Also, occupational findings for males indicate a linear trend for each disorder. Furthermore, the scatter diagram for occupational status reveals that,

with the exception of a few deviant cases, the relationship is gradual with a generally consistent decline in rates of brain disorder and schizophrenia as occupational status increases. It is not likely that the effect of mental illness on class status would have such a constant effect at all levels of the class structure.[10]

Note that occupations in the middle and upper range of socioeconomic values require more than simple occupational skills; it seems most doubtful that differences between occupations in this range would be due to social

TABLE 4

CRUDE HOSPITALIZATION RATES PER 100,000 POPULATION BY AGE, EDUCATION, AND DIAGNOSIS—25- TO 64-YEAR-OLD MALES AND FEMALES, TENNESSEE, 1956–65

	EDUCATION					
AGE AND DIAGNOSIS	0–4	5–7	8	9–11	12	+12
25–29:						
Brain disorders	146	87	66	45	30	11
Schizophrenia	428	398	285	295	173	134
Other	173	269	266	392	275	152
30–34:						
Brain disorders	166	100	92	55	20	14
Schizophrenia	326	390	361	277	204	150
Other	171	276	341	341	236	160
35–39:						
Brain disorders	211	131	88	57	37	23
Schizophrenia	258	247	225	250	151	128
Other	249	274	290	276	238	177
40–44:						
Brain disorders	189	100	72	66	51	51
Schizophrenia	343	215	223	120	125	85
Other	250	295	315	284	284	255
45–49:						
Brain disorders	160	112	91	118	97	32
Schizophrenia	182	147	114	107	99	85
Other	280	235	401	352	348	207
50–54:						
Brain disorders	250	171	147	91	122	58
Schizophrenia	86	119	118	55	67	42
Other	263	270	364	375	347	203
55–59:						
Brain disorders	263	234	284	192	172	85
Schizophrenia	63	44	64	57	80	30
Other	206	232	295	279	322	231
60–64:						
Brain disorders	429	330	439	228	417	194
Schizophrenia	41	21	29	23	43	28
Other	95	144	215	154	238	175

SOURCE.—Population base: U S. Bureau of the Census (1963 315–16).

[10] According to Dunham (1965, p. 191), a drift effect for schizophrenia would probably be most apparent if there were an unusually high schizophrenic rate for the lowest stratum, with little or no difference between positions in the middle and upper range of socioeconomic values.

selection based on "mental health." For example, although teachers have a higher socioeconomic score than electricians, the skills of the latter are sufficiently high so that persons who are incapable of performing as electricians for psychiatric reasons are not likely to be capable of functioning as teachers. Consequently, the difference in rates of brain disorder (39 vs. 231) is probably not due solely to drift. (Other comparisons may be made from table 1 for both brain disorders and schizophrenia [e.g., rate of 68 for salespersons and clerks vs. rate of 390 for construction and maintenance painters].) It would seem, then, that without rejecting the role of drift altogether, there is something about socioeconomic status that contributes to schizophrenia (Kohn 1976) and to mental disorders with an organic base.

Socioeconomic Stress

Most interpretations of the role of socioeconomic factors in mental illness emphasize greater stress in the lower socioeconomic strata stemming from such things as blocked aspirations and status frustration. Obviously such frameworks are not valid for brain disorders since these disorders have a physical basis and are not due to socially induced psychological stress.

They also do not provide very plausible accounts for schizophrenia. The most coherent and explicit conceptual framework which emphasizes stress is Kohn's formulation (1972*a*, 1972*b*, 1976), which states that genetic vulnerability together with stress and limited-rigid conceptions of reality (both of which Kohn believes are more prevalent among lower-class individuals), combine to produce a high rate of schizophrenia in the lower class. The formulation seems to assume that genetic vulnerability, stress, and the impaired ability to deal with problematic and stressful situations are all higher for members of the lower socioeconomic strata"[11] (Kohn 1976, p. 180). Unfortunately for this framework, there is little evidence that lower-status individuals suffer from more situationally induced stress or that they are not as well equipped to deal with it when stressful events occur.[12] Also, while

[11] Actually, the formulation need not necessarily stipulate that all three variables are higher in the lower socioeconomic strata There may be statistical interactions among the three factors and class status so that forces generating schizophrenia may be greater in the lower strata even though the values of each contributing factor are not.

[12] The literature on stress, especially as related to socioeconomic status, is ambiguous at best. This is partly due to the definition and measurement of stress and the designation of external events which are stress inducing. For example, the Dohrenwends view stress as the consequence of "stressors," defined as "objective events that disrupt or threaten to disrupt the individual's usual activities" (1973, p. 115). One may wonder whether disruption of usual activities produces stress any more frequently than the absence of disruption. In any case, the Dohrenwends conclude that "there is no firm evidence that the overall rate of stressors varies with social class" (ibid., p. 119). At the same time, they apparently believe that lower-class persons are less able to cope with stress, either because of external factors (e.g., support from family) or internal

there is some limited evidence to suggest that life changes and problems in living are precipitating factors in schizophrenia (Rogler and Hollingshead 1965; Brown and Birley 1968; Birley and Brown 1970), the evidence for stress itself is far from compelling (Mechanic 1972). One review indicates that the evidence is too unsatisfactory to permit definitive conclusions about a relationship between stress and mental illness (Jaco 1973). More significantly, there appears to be no evidence showing that stress associated with schizophrenia is linked to socioeconomic status.[13] In comparison, there is substantial evidence that a number of biological and physiological factors, which would appear to be related to socioeconomic status, are most definitely involved in schizophrenic disorders.[14]

Labeling

It might be argued that the present findings are due to differential societal responses to individuals of different socioeconomic status who do not really differ in mental status. Some evidence does indicate that responses of mental health professionals may be influenced by the socioeconomic status of patients (Gardner and Badigan 1966; Fried 1969, p. 115; Michael 1967). It would seem, however, that this would make for a stronger relationship for

factors (e.g , lower intelligence) Hence, although they recognize serious limitations on the quality of the data, they suggest that the best hypothesis is that severity of stress is greater in the lower class than in the middle class (ibid., p. 132) We are not convinced. This is due partly to the definition of stressor, partly to the authors' assumptions as to how various factors (e g , intelligence) mediate stress, and partly to the questionable effects (which the Dohrenwends recognize) that certain attitudinal factors have in mediating the effects of stressors Finally, there is no actual evidence that stress is higher in one class than in another We are also inclined to believe that the adaptive and coping ability of the socioeconomically disadvantaged is far greater than many suppose it to be. People learn to cope by developing skills to handle difficult life situations, and, as David Mechanic notes, the disadvantaged have more opportunity to develop such skills than persons more fortunate do (1972, p. 307). Glen Elder's (1974) study of the children of the depression is particularly relevant to this point. We believe that the only definitive conclusion one can make about socioeconomic status and stress is that the relationship is either nonexistent or the evidence is too limited and ambiguous to permit a conclusion.

[13] For example, although Kohn (1976) cites Rogler and Hollingshead's study (1965) as supporting evidence for his stress formulation, this study produces no evidence showing that rates of schizophrenia are higher in the lower class because of stress This study does not give rates by class, since it was limited to lower-class subjects—a sample of lower-class schizophrenics and a sample of matched controls Our point is not that Kohn fails to recognize this, since he does (e.g., 1972*a*, p. 298) Rather, it is to note that even if psychic stress is related to schizophrenia there is no evidence that such stress is linked to socioeconomic status (on this point, see especially Brown and Birley 1968, p. 212)

[14] This evidence would not be inconsistent, of course, with physiological and biochemical models of stress (e.g., Selye 1956). To the extent that mental illness has a biological base, it could be a consequence of stress on the physiological and biochemical systems.

the other disorders. There is probably less class bias, or error from any source, in organic than in functional diagnoses overall,[15] and evidence would indicate that interjudge reliability of schizophrenia diagnosis is reasonably high (Kety 1974, p. 960).

IMPLICATIONS

One implication of our analysis is that the results of studies which have reported a relationship between socioeconomic status and mental illness may be primarily due to the inclusion of patients with brain and schizophrenic disorders, or their precursors, in the analysis. If our interpretation is valid, the overall relationship is due more to the role of medical rather than psychiatric and psychological factors. A relationship of socioeconomic status to mental illness is thus consistent with the medical model.

At the same time, the explanation we offer is basically sociological. The various factors which have been postulated as significant are aspects of social locations in society that differentiate between groups and categories of persons. In a very general sense they may be viewed as components of class "subcultures," referring as they do to a range of environmental conditions, behavior patterns and attitudes, and access to medical care which distinguish among socioeconomic strata in society.[16] The interpretation thus emphasizes the role of socioeconomic factors in creating mental illness either through direct action on physical states or in mediating the effects of physical conditions which are conducive to the production of mental illness. The formulation would appear to be specific to brain and schizophrenic disorders since data do not indicate that other disorders (manic-depression, neuroses, and personality disorders) are associated with socioeconomic status (but

[15] Torrey states that "the signs and symptoms of possible brain disease can be taught to a nondoctor in one or two weeks . . ." (1974, p. 67).

[16] It is also conceivable that findings for both brain and schizophrenic disorders reflect a tendency for higher status persons to seek care in the private sector. This, however, would not explain why such a tendency does not exist among persons in the "other" category. Also, in most instances private psychiatric care requires substantial economic resources which are associated with and largely determined by educational achievement as well as verbal and intellectual capacities which also come largely as a result of education. It is likely that the most significant distinction between individuals in both these respects is between the college educated and the noncollege educated. Results indicate, however, that these are not the places where significant breaks occur in the relationships for brain disorders and schizophrenia (see figure 1); overall, the relationships are essentially linear. In addition, a negative relationship with socioeconomic status has been reported in studies even when private hospital patients have been included (e.g., Hollingshead and Redlich 1958; Dunham 1965). Also, Dunham reports that almost all patients of private practitioners spent time as hospital inpatients (1965, p. 262); and Pasamanick, Scarpiti, and Dinitz (1967, p. 138) find that the socioeconomic distribution for state hospital schizophrenics is similar to that for schizophrenics referred to a treatment center by private physicians and welfare and other agencies.

see n. 4 above), and there is less reason to believe that biological factors are important in most of these other disorders (but see n. 8 above).

There is a policy implication here. To the extent that differences in brain disorders and schizophrenia are due to differences in access to quality medical care, government policy which reduces class differences in medical care would reduce class differences in mental illness. This would be the case even if differences in access to mental health professionals were unchanged. According to our formulation, factors which are involved in the class-related etiology of mental illness have an effect on mental states because they first have either a direct or a mediating effect on physical states. Hence, programs which remove economic barriers to quality (nonpsychiatric) medical care should reduce class differences in mental illness as well as in mortality and other forms of morbidity. Also, the elimination of occupational hazards would probably reduce the rate of brain disorders for low-status occupations.[17]

Finally, in concluding, we would make three points.

First, our focus on an integrated sociomedical model interpretation does not altogether deny the validity of various views which contend that the medical model is inapplicable for understanding mental disorders. Our findings are for only one variable among many. "Problems in living," labeling, conflicting expectations, crowded and urban living, even stress—all may be factors in mental illness, and, indeed, they may be more important than socioeconomic status. Our effort has not been to provide an overall theory of mental illness/hospitalization, or even of brain and schizophrenic disorders.

Second, our formulation stipulates that the relevant socioeconomic factors are numerous. There is no *one* socioeconomic factor that is primary. There are many specific factors,[18] most of which can be classified in the categories of physical environment, attitudes and behavior, medical care, and exposure to infectious agents. Therefore, by "integrated sociomedical model" we mean that various aspects of the social environment have direct or mediating effects on biological conditions which in turn have physiological and organic

[17] We note, however, that a policy implication is not the same as a policy recommendation. Whether the above policies are to be recommended would depend on how much they would cost in comparison with the benefits they would generate. This is difficult to assess because, while the costs of such programs can be estimated with a reasonable degree of accuracy (at least in economic terms), the benefits cannot (e.g., the number of cases of mental illness that would be prevented is almost impossible to estimate).

[18] In a similar vein, Fried states: "It may be . . that our usual expectation that a single major factor will account for a large part of the variance is inappropriate for the investigation of psychiatric disorder, that many diverse factors have minor and independent determining effects and that only their cumulative effect serves as a potent and highly general determinant of mental illness" (1969, p. 163). In contrast, however, Fried's focus, as most others, is on intervening psychological states (1969, pp. 137–55, especially at 155—"one can only conclude that if poverty per se is closely linked to psychiatric disorder, it must be through some psychological mechanism . .").

consequences that can be understood only in terms of a medical model.

Third, although our results are consistent with the bulk of the evidence, the proportion of brain and schizophrenic patients in our study is probably higher than in some studies, especially community surveys. At the same time, community surveys may actually underreport the number of physically based mental disorders. The Midtown Manhattan study reports a negative correlation between socioeconomic status and reports of poor physical health as well as between socioeconomic status and psychiatric symptoms. A positive relationship between reports of physical illness and rates of psychiatric disorder thought to be psychogenic is also indicated. Interpretation of these findings is problematic, of course, because psychophysiological symptoms form an important basis for psychiatric assessments (for a discussion of this problem, see Crandall and Dohrenwend 1967, pp. 1527–37). But, the psychiatric conditions may be due to the physical condition rather than vice versa. Indeed, Derogatis, Yeveseroff, and Wittlesberger believe their findings indicate that the relationship between socioeconomic status and psychiatric disorder may represent an extension of the relationship for physical disease (1975, p. 190).

Symptoms such as trembling hands and heart palpitations are treated in mental-health scales as physiological expressions of mental unease. It is at least as plausible to consider such symptoms as reflections of physical conditions. (In some instances, of course, they may be contributing factors to mental illness.) Thus, the lower socioeconomic strata may show a higher prevalence of mental illness in community studies only when mental illness is measured in terms of the number of psychophysiological symptoms. The position gains further credence from studies of socioeconomic status and mental illness that are based on the Langner scale, which may be weighted in favor of physiological symptoms (see Muller 1971, p. 601). Such findings may be confounded by either socioeconomic reactions to physical illness or the actual manifestation thereof (see Crandall and Dohrenwend 1967, p. 1535). In either case, data would indicate that the relationship between socioeconomic status and mental illness is, at least in part and perhaps for the most part, an aspect of a more general relationship between socioeconomic status and physical status.

REFERENCES

Allen, M. G. 1976. "Twin Studies of Affective Illness." *Archives of General Psychiatry* 33 (December): 1476–78.

Alpert, Joel J., John Kosa, Robert J. Haggerty, Leon Robertson, and Margaret C. Haggerty. 1969. "The Types of Families That Use an Emergency Clinic." *Medical Care* 7 (January–February): 55–61.

Anderson, Ronald, and Odin W. Anderson. 1977. "Trends in the Use of Health Services." Pp. 386–407 in *Handbook of Medical Sociology*, edited by Howard E. Freeman, Sol Levine, and Leo G. Reeder. 3d ed. Englewood Cliffs, N.J.: Prentice-Hall.

Antonovsky, Aaron. 1972. "Social Class, Life Expectancy, and Overall Mortality." Pp. 5–30 in *Patients, Physicians and Illness*, edited by E. Gartley Jaco New York: Free Press.

Banfield, Edward C. 1974. *The Heavenly City Revisited* Boston· Little, Brown

Berkanovic, Emil, and Leo G. Reeder. 1974 "Can Money Buy the Appropriate Use of Services? Some Notes on the Meaning of Utilization Data." *Journal of Health and Social Behavior* 14 (June): 93–99

Bice, Thomas L., Robert Eichhorn, and P D. Fox 1972. "Socioeconomic Status and Use of Physician Services." *Medical Care* 3 (May–June)· 261–71

Birley, J. L. T , and G W. Brown. 1970. "Crises and Life Changes Preceding the Onset or Relapse of Acute Schizophrenia Clinical Aspects " *British Journal of Psychiatry* 116 (January): 327–33.

Boklage, Charles E. 1977. "Schizophrenia, Brain Asymmetry and Twinning· Cellular Relationship with Etiological and Possibly Prognostic Implications " *Biological Psychiatry* 12 (February): 19–35.

Brackbill, G. A , and H. J Fine 1956. "Schizophrenia and Central Nervous System Pathology." *Journal of Abnormal and Social Psychology* 52 (May): 310–13

Bremer, Johan 1951. "A Social Psychiatric Investigation of a Small Community in Northern Norway." *Acta Psychiatrica et Neurologica* Supplementum 62

Bronfenbrenner, Urie. 1958. "Socialization and Social Class through Time and Space." Pp. 400–425 in *Readings in Social Psychology*, edited by Eleanor E. Maccoby, Theodore M. Newcomb, and E L Hartley 3d ed. New York: Holt, Rinehart & Winston.

Brown, George W , and J L. T Birley 1968 "Crisis and Life Changes and the Onset of Schizophrenia." *Journal of Health and Social Behavior* 9 (September)· 203–14

Brune, Gunter B. 1971. "The Somatically Determined Psychoses." Pp 43–78 in *Biochemistry, Schizophrenia, and Affective Illnesses*, edited by H. Himwich. Williams & Wilkins.

Campion, Edward, and Gary Tucker 1973. "A Note on Twin Studies, Schizophrenia and Neurological Impairment." *Archives of General Psychiatry* 20 (October)· 461–64

Committee on Nomenclature and Statistics of the American Psychiatric Association 1952. *Diagnostic and Statistical Manual, Mental Disorders* Washington, D.C : American Psychiatric Association.

Crandall, Dewitt L., and Bruce P Dohrenwend 1967 "Social Relationships among Psychiatric Symptoms, Organic Illness and Social Class." *American Journal of Psychiatry* 123 (June): 1527–37.

Davidson, K., and C. R. Bagley. 1969. "Schizophrenia-like Psychoses Associated with Organic Disorders of the Central Nervous System: A Review of the Literature " Pp. 113–84 in *Current Problems in Neuropsychiatry* (*British Journal of Psychiatry* Special Publication no. 4), edited by R N. Herrington Ashford, England Headley Brothers

Department of Health, Education, and Welfare. 1972. *Physicians' Visits: Volume and Interval since Last Visit, United States, 1969*. Series 10, no 75 Washington, D C · Government Printing Office.

Derogatis, Leonard, H. Yeveseroff, and B. Wittlesberger. 1975. "Social Class, Psychological Disorder and the Nature of the Psychopathologic Indicator " *Journal of Consulting and Clinical Psychology* 43 (April): 183–91

Dohrenwend, Barbara Snell, and Bruce P. Dohrenwend. 1973. "Class and Race as Status Related Sources of Stress." Pp. 111–40 in *Social Stress*, edited by Sol Levine and Norman A. Scotch. Chicago: Aldine.

Dohrenwend, Bruce P., and Barbara Snell Dohrenwend 1969 *Social Status and Psychological Disorder: A Causal Inquiry*. New York: Wiley

———. 1974. "Social and Cultural Influences on Psychopathology " *Annual Review of Psychology* 25:417–52.

———. 1977. "Sex Differences and Psychiatric Disorders " *American Journal of Sociology* 81 (May): 1447–54.

Duff, Robert S., and August B. Hollingshead. 1968. *Sickness and Society*. New York: Harper & Row.

Duncan, Otis Dudley. 1961. "A Socio-economic Index for all Occupations," and "Properties and Characteristics of the Socioeconomic Index." Pp. 109–61 in *Occupations and Social Status,* edited by Albert J. Reiss, Jr. New York: Free Press.
Dunham, H. Warren. 1965. *Community and Schizophrenia.* Detroit: Wayne State University Press.
Elder, Glen. 1974. *Children of the Great Depression.* Chicago: University of Chicago Press.
Erlanger, Howard S. 1974. "Social Class Differences in Parents' Uses of Physical Punishment." Pp. 150–58 in *Violence in the Family,* edited by Suzanne K Steinmetz and Murray A. Strauss. New York: Dodd, Mead.
Fried, Marc. 1969. "Social Differences in Mental Health." Pp 113–67 in *Poverty and Health: A Sociological Analysis,* edited by J Kosa, A. Antonovsky, and I. K. Zola. Cambridge, Mass.: Harvard University Press.
Gardner, E. A., and H. M. Badigan. 1966. "A Longitudinal Comparison of Psychiatric Service." *American Journal of Orthopsychiatry* 36 (May): 818–28.
Gershon, E. S., D. L. Dunner, and F. K. Goodwin 1971. "Toward a Biology of Affective Disorders." *Archives of General Psychiatry* 25 (July)· 1–15
Gil, David G 1970 *Violence against Children.* Cambridge, Mass.: Harvard University Press.
Gillis, L. S , J. B. Lewis, and M. Slabbert. 1968. "Psychiatric Disorder amongst the Coloured People of the Cape Peninsula: An Epidemiological Study." *British Journal of Psychiatry* 114 (December) 1575–87.
Goffman, Erving. 1961. *Asylums: Essays on the Social Situation of Mental Patients.* Garden City, N.Y.: Anchor
——— 1971. *The Insanity of Place.* London: Allen Lane
Gold, Martin 1958. "Suicide, Homicide, and the Socialization of Aggression " *American Journal of Sociology* 43 (May)· 651–61
Goode, David J , Herbert Y. Meltzer, John W. Crayton, and Terry Mazura. 1977. "Physiological Abnormalities of the Neuromuscular System in Schizophrenia." *Schizophrenia Bulletin* 3 (1): 121–38.
Goode, William J. 1971. "Force and Violence in the Family." *Journal of Marriage and the Family* 33 (November)· 624–36.
Gottesman, I., and I. Shields. 1972. *Schizophrenia and Genetics: A Twin-Study Vantage Point.* New York: Academic Press.
Gove, Walter R 1972 "The Relationship between Sex Roles, Mental Illness and Marital Status." *Social Forces* 51 (September): 34–44.
———. 1976. "Reply to Imersheim and Simons (1976) and Scheff (1975)." *American Sociological Review* 41 (June): 564–67.
Gove, Walter R., and Jeanette F. Tudor. 1973. "Adult Sex Roles and Mental Illness." *American Journal of Sociology* 73 (January): 812–35.
Guralnick, Lillian. 1963*a* *Mortality by Occupation and Cause of Death among Men 20 to 64 Years of Age: United States, 1950.* Washington, D.C.: U.S. Department of Health, Education, and Welfare.
———. 1963*b*. *Mortality by Industry and Cause of Death among Men 20 to 64 Years of Age: United States, 1950.* Washington, D.C.: U.S. Department of Health, Education, and Welfare.
Helgason, Tomas. 1964. "Epidemiology of Mental Disorders in Iceland." *Acta Psychiatrica Scandinavica.* Supplementum 173.
Hendrick, I. 1928. "Encephalitis Lethargica and the Interpretation of Mental Disease." *American Journal of Psychiatry* 84 (May): 989–1014.
Herman, W. 1972. "The Poor: Their Needs and Health Services Available to Them." *Annals* 399 (January): 12.
Hollingshead, August, and Frederick Redlich. 1958. *Social Class and Mental Illness.* New York: Wiley.
Hyde, Robert W., and Lowell V. Kingsley. 1944. "The Relation of Mental Disorders to the Community Socioeconomic Level." *New England Journal of Medicine* 231 (October): 543–48.

Jaco, E. Gartley. 1960. *The Social Epidemiology of Mental Disorders.* New York: Russell Sage.

———. 1973. "Mental Illness in Response to Stress." Pp 210–27 in *Social Stress,* edited by Sol Levine and Norman A Scotch Chicago: Aldine.

Kety, Seymour S. 1974. "From Rationalization to Reason" *American Journal of Psychiatry* 131 (September): 957–63

Klerman, Gerald L. 1975. "The Clinical Relevance of Recent Research for Treatment and Prevention of Schizophrenia." Pp. 80–101 in *Schizophrenia: Biological and Psychological Perspectives,* edited by G. Usdin New York: Brunner/Mazel.

Kohn, Melvin L. 1968. "Social Class and Schizophrenia A Critical Review." Pp 155–73 in *The Transmission of Schizophrenia,* edited by David Rosenthal and Seymour S. Kety. Oxford: Pergamon.

———. 1969. *Class and Conformity: A Study in Values* Homewood, Ill : Dorsey.

———. 1972*a*. "Class, Family, and Schizophrenia: A Reformulation." *Social Forces* 50 (March): 295–304.

———. 1972*b*. "Rejoinder to David Mechanic." *Social Forces* 50 (March): 310–13.

———. 1976. "The Interaction of Social Class and Other Factors in the Etiology of Schizophrenia." *American Journal of Psychiatry* 133 (February): 177–80.

Koos, Earl. 1954. *The Health of Regionsville: What the People Thought and Did about It.* New York: Columbia University Press.

Langner, T. S , and S. T. Michael. 1963. *Life Stress and Mental Health.* New York: Free Press.

Lerner, Monroe 1969 "Social Differences in Physical Health." Pp. 69–112 in *Poverty and Health: A Sociological Analysis,* edited by J Kosa, A Antonovsky, and I. K. Zola. Cambridge, Mass : Harvard University Press

Levinger, George. 1966. "Source of Marital Satisfaction among Applicants for Divorce " *American Journal of Orthopsychiatry* 36 (October): 804–6

Lin, Tsung-yi. 1955. "A Study of the Incidence of Mental Disorder in Chinese and Other Cultures." *Psychiatry* 16 (November): 313–36.

Mechanic, David 1968. *Medical Sociology.* New York: Free Press

———. 1972. "Social Class and Schizophrenia. Some Requirements for a Plausible Theory of Social Influence." *Social Forces* 50 (March): 305–9.

Michael, S. T. 1967. "Social Class and Psychiatric Treatment." *Journal of Psychiatric Research* 5 (May): 243–54.

Miles, David L. 1977. *Health Care Evaluation Project: Terminal Progress Report, July 1977.* Rockville, Md.: National Center for Health Services Research.

Miller, Walter. 1958. "Lower Class Culture as a Generating Milieu of Gang Delinquency." *Journal of Social Issues* 14 (3): 5–19.

Morgan, David. 1975 "Explaining Mental Illness." *Archives Europeennes de Sociologie* 16 (November): 262–80

Muller, David L. 1971. "Discussion." *American Journal of Psychiatry* 128 (November): 600–603.

Pasamanick, Benjamin, Frank R. Scarpiti, and Simon Dinitz. 1967. *Schizophrenia in the Community: An Experimental Study in the Prevention of Hospitalization.* New York: Appleton.

Remington, F. B., and S. L. Rubert. 1962. "Why Patients with Brain Tumors Come to a Psychiatric Hospital." *American Journal of Psychiatry* 112 (February): 256–63.

Richardson, William C. 1972. "Poverty, Illness, and the Use of Health Services in the United States." Pp. 240–49 in *Patients, Physicians and Illness,* edited by E. Gartley Jaco. New York. Free Press.

Rogler, Lloyd D., and August B. Hollingshead. 1965. *Trapped: Families and Schizophrenia.* New York: Wiley.

Rosenstock, Irwin. 1969. "Prevention of Illness and Maintenance of Health." Pp. 168–90 in *Poverty and Health: A Sociological Analysis,* edited by J. Kosa, A. Antonovsky, and I. K. Zola. Cambridge, Mass.: Harvard University Press.

Ross, John A. 1962. "Social Class and Medical Care." *Journal of Health and Social Behavior* 3 (Spring): 35–40.

Roth, J. A. 1969. "The Treatment of the Sick." Pp. 214–43 in *Poverty and Health: A Sociological Analysis,* edited by J. Kosa, A. Antonovsky, and I. K. Zola. Cambridge, Mass.: Harvard University Press.

Rushing, William A., and David L. Miles. 1977. "Physicians, Physicians' Assistants, and the Social Characteristics of Patients in Southern Appalachia." *Medical Care* 15 (December): 1004–13.

Sanders, E. P. 1964. "Relation between Accident Incidence and Type and Level of Job." *Psychological Reports* 14 (June): 670.

Scheff, Thomas J. 1966. *Being Mentally Ill.* Chicago: Aldine.

———. 1974. "The Labelling Theory of Mental Illness." *American Sociological Review* 39 (June): 444–52.

Selye, Hans. 1956. *The Stress of Life.* New York: McGraw-Hill.

Shields, James. 1968. "Summary of the Genetic Evidence." Pp. 95–125 in *The Transmission of Schizophrenia,* edited by David Rosenthal and Seymour S. Kety. London: Pergamon.

Slater, Eliot, A. W. Beard, and E. Glithero. 1963. "The Schizophrenia-like Psychoses of Epilepsy." *British Journal of Psychiatry* 109 (January): 95–105.

Steinmetz, Suzanne K. 1974. "Occupational Environment in Relation to Physical Punishment and Dogmatism." Pp. 166–72 in *Violence in the Family,* edited by Suzanne K. Steinmetz and Murray A. Straus. New York: Dodd, Mead.

Straus, Murray. 1971 "Some Social Antecedents of Physical Punishment: A Linkage Theory Interpretation." *Journal of Marriage and the Family* 33 (November): 658–63.

Studnicki, J., R. M. Saywell, and W. Wiecheter. 1976. "Foreign Medical Graduates and Maryland Medicaid." *New England Journal of Medicine* 294 (May): 153–57.

Symthies, J. R. 1975. "Biochemical Abnormalities Associated with Schizophrenia." *British Journal of Psychiatry,* Special Issue, 9: 41–47.

Szasz, Thomas S. 1961. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct.* New York: Dell.

Taylor, David C. 1975. "Factors Influencing the Occurrence of Schizophrenia-like Psychosis in Patients with Temporal Lobe Epilepsy." *Psychological Medicine* 5 (August): 249–54.

———. 1977. "Epilepsy and the Sinister Side of Schizophrenia." *Developmental Medicine and Child Neurology* 19 (June): 403–7.

Toby, Jackson. 1966. "Violence and the Masculine Ideal: Some Qualitative Data." *Annals of the American Academy of Political and Social Science* 364 (March): 20–27.

Torrey, E. Fuller. 1974. *The Death of Psychiatry.* Radnor, Pa.: Chilton.

Torrey, E. Fuller, and Michael R. Peterson. 1976. "The Vital Hypothesis of Schizophrenia." *Schizophrenia Bulletin* 2 (1); 136–45.

Trimbel, Michael. 1977. "The Relationship between Epilepsy and Schizophrenia: A Biochemical Hypothesis." *Biological Psychiatry* 12 (February): 299–304.

U.S. Bureau of the Census. 1963. *U.S. Population: 1960.* Vol. 1. *Characteristics of the Population.* Part 44, Tennessee. Washington, D.C.: Government Printing Office.

Winslow, Robert W. 1968. *Crime in a Free Society: Selections from the President's Commission on Law Enforcement and Administration of Justice.* Belmont, Calif.: Dickenson.

The Community Question: The Intimate Networks of East Yorkers[1]

Barry Wellman

University of Toronto and the Netherlands Institute for Advanced Study

The Community Question has set the agenda for much of sociology. It is the question of how large-scale social systemic divisions of labor affect the organization and content of primary ties. Network analysis is proposed as a useful approach to the Community Question, because, by focusing on linkages, it avoids the a priori confinement of analysis to solidary groupings and territorial units. Three contentions about the Question are evaluated: arguments that Community is Lost, Saved, or Liberated. Data are presented about the structure and use of the "intimate" networks of 845 adult residents of East York, Toronto. Intimate networks are found to be prevalent, composed of both kin and nonkin, nonlocal, asymmetric, and of sparse density. Help in dealing with both emergencies and everyday matters is available from almost all intimate networks, but from only a minority of intimate ties. The data provide broad support for the Liberated argument, in conjunction with some portions of the Saved argument.

COMMUNITY AS NETWORK

The Community Question has set the agenda for much of sociology. It is the question of how large-scale social systemic divisions of labor affect the organization and content of primary ties. The Question thus has formed a crucial sociological nexus between macroscopic and microscopic analysis. It has posed the problem of the structural integration of a social system and the interpersonal means by which its members have access to scarce resources.

In considering the Community Question, sociologists have been espe-

[1] This paper has been revised a number of times, and I am grateful to the following people who have commented extensively on some form of it: S. D. Berkowitz, Y. Michal Bodemann, L. S. Bourne, Ronald Burt, Bonnie Erickson, Linton Freeman, Harriet Friedmann, Joseph Galaskiewicz, Leslie Howard, Nancy Howell, Edward Lee, Barry Leighton, J. Clyde Mitchell, Livianna Mostacci, Walter Phillips, Chris Pickvance, Norman Shulman, Charles Tilly, Jack Wayne, Beverly Wellman, Harrison White, and anonymous *AJS* referees. The following agencies have supported portions of the research: Canada Council, Canada Ministry of Manpower and Immigration, Clarke Institute of Psychiatry, Laidlaw Foundation, Netherlands Institute for Advanced Study, Ontario Ministry of Health, Social Science and Humanities Research Council of Canada, and Urban Housing Markets Program (Centre for Urban and Community Studies, University of Toronto).

cially concerned with assessing the impact of industrialization and bureaucratization on a variety of primary ties: in the neighborhood, in kinship groups, in interest groups, and on the job. Urban sociologists, in particular, have been interested in this matter. From Tönnies ([1887] 1955) to Fischer (1977), they have investigated the effects of industrial bureaucratic social systems on communal structures, with particular reference to the following: (*a*) the increased scale of the nation-state's activities, with a concomitant low level of local community autonomy and solidarity (e.g., Tilly 1973, 1975); (*b*) the development of narrowly instrumental bureaucratic institutions for production and reproduction (cf. Castells 1976); (*c*) the large size of cities, with the consequent population and organizational potential for diverse interest groups; (*d*) the high social density of interactions among the segments of the population (even where spatial density is low), with the ensuing complexities of organizational and ecological sorting;[2] (*e*) the diversity of persons with whom city dwellers can come into contact under conditions of heightened mobility; and (*f*) widespread networks of cheap and efficient transportation facilities, letting contact be maintained with greater ease and over longer distances (cf. Meier 1968). The increased velocity of transactions facilitates interactional density: the large-scale city is accessible, centralized control can more effectively be imposed, and links to diverse social networks can more readily be maintained.

Unfortunately, in many community analyses, the basic structural concerns of the Community Question have become confounded with two other sociological concerns: (1) a preoccupation with the conditions under which solidary sentiments can be maintained, reflecting a continuing, overarching sociological concern with normative integration and consensus; and (2) a preoccupation with locating primary ties in local areas, reflecting urban sociology's particular concern with spatial distributions.

As a result of this confounding, the fundamentally structural Community Question has often been transmuted into a search for local solidarity, rather than a search for functioning primary ties, wherever located and however solidary. (It is my underlying argument that the proper concern of sociologists is the analysis of social structure and social linkages, with questions of social sentiments and spatial distribution holding important, but secondary, positions.) Analyses have tended to take as their starting point extrinsic mappings of local area boundaries and then proceeded to enquire into the extent of communal interaction and sentiment within these

2 Early formulations of this point were in terms of high spatial density, but such formulations have been called into question by both the suburban dispersion and doubts about the social effects of crowding and density. In any event, analyses of spatial density tend to use interactional density as an intervening variable (see Freedman 1975), and if the questionable premise is not valid, the useful conclusion still remains (see Abu-Lughod 1969; Tilly 1970).

boundaries. They have thus assumed, a priori, that a significant portion of an urbanite's primary ties are organized by locality. Such a territorial perspective, searching for answers to the Community Question only within bounded population aggregates, has been especially sensitive to the evaluation of community solidarity in terms of shared values (see the discussions in Friedmann 1974; Howard 1974). Consequently, when there has been an observed dearth of locally organized solidary behavior and sentiments, the assumption has easily been made that "community" has decayed. Such assumed losses of community have been prevalent in the contemporary milieu of frequent residential mobility and spatially dispersed networks and activities.

Conceptualizing the interpersonal life of the city dweller as the central node linking together complex network structures leads to quite different analytic concerns from conceptualizing it as a membership in a discrete solidarity. Hence I propose an examination of the Community Question from a network analytic perspective. The utility of the network perspective is that it does not take as its starting point putative solidarities—local or kin—nor does it seek primarily to find and explain the persistence of solidary sentiments. It attempts to avoid individual-as-unit research perspectives, with their inherent social psychologistic explanatory bases, seeing internalized attitudes as determining social relations.

Instead, social network analysis is principally concerned with delineating structures of relationships and flows of activities. By looking directly at linkages rather than at solidarities, the network perspective enables us to focus directly on the basic structural issues posed by the Community Question.[3] Such an approach can do much to free the study of community from normative and spatial predilections.

This paper presents a social network analysis of the Community Question debate in urban sociology, as informed by a study of close ("intimate") ties in East York, Toronto. I first review three current Community Question arguments from a network analytic perspective: contentions that

[3] Although often mistakenly thought of as a collection of techniques, network analysis is essentially an analytic perspective which focuses on structured relationships between individuals and collectivities. As yet, there is no commonly agreed upon definition. Some of the salient characteristics of network analysis are that it gives attention to (*a*) structured patterns of relationships and not the aggregated characteristics of individual units analyzed without reference to their connectivity; (*b*) complex network structures and not just dyadic ties; (*c*) the allocation of scarce resources through concrete systems of power, dependency, and coordination; (*d*) network boundaries, clusters, and cross-linkages; and (*e*) complex structures of reciprocal relationships and not just symmetrical relationships or simple hierarchies. For summaries of the network perspective see Emerson (1962), White (1965), Mitchell (1969, 1974), Barnes (1972), Kemper (1972), Craven and Wellman (1973), and White, Boorman, and Breiger (1976). See also the bibliographies of Wellman and Whitaker (1974), Freeman (1976), and Klovdahl (1977).

Community is now "Lost," "Saved," or "Liberated" (the arguments are more fully discussed in Wellman and Leighton 1979). I then examine these arguments in the light of the data. Last, I discuss some of the broader implications of this review and the findings for the analysis of the Community Question.

COMMUNITY: LOST, SAVED, LIBERATED

Community Lost

The Lost argument was the first urban sociological response to the Community Question (e.g., Tönnies [1887] 1955), and it is still significantly influencing the current debate.[4] The argument holds many urban phenomena to be concrete and concentrated manifestations of industrial bureaucratic societies. It contends that the division of labor in these societies has attenuated communal solidarities. Primary relationships in the city now are "impersonal, transitory and segmental" (Wirth 1938, p. 12). Instead of being fully incorporated into a single solidary community, urbanites are seen as being limited members of multiple social networks, sparsely knit and loosely bounded. Their weak, narrowly defined, and disorganized ties are rarely available or useful for help in dealing with contingencies. Consequently, urbanites are now bound to the city by webs of secondary affiliations.

The Lost argument has occupied an important place in North American thought, from Jeffersonian antiurbanism through Progressive reformism (e.g., Woodsworth [1911] 1972) and "Chicago school" urban sociology (e.g., Park 1925*a*; Wirth 1938) to recent jeremiads against "mass society," both scholarly (e.g., Nisbet 1969) and popular (e.g., *Death Wish* [1974]). The argument's emphasis on the alleged disorganizing effects of attenuated communal solidarities has been reflected in substantive accounts of such diverse areas as collective action, crime, migration, poverty, and suburbia (see the critical reviews of Valentine 1968; Feagin 1973; Mostacci 1976).[5]

The Lost argument has usefully sharpened awareness of potential relationships between industrial bureaucratic divisions of labor and structures of primary ties. Yet, because of its assumption that strong primary ties naturally occur only in densely knit, self-contained solidarities, the argument has unduly neglected the question of whether primary ties have

[4] See the reviews of Stein (1960), Nisbet (1969), Gusfield (1975), and Castells (1976). Good examples of nonurban sociological uses of the Lost argument can be found in the political analyses of Kornhauser (e.g., 1968) and Gurr (e.g., 1969); see also Tilly's critiques (e.g., 1978).

[5] See White and White's (1962) and Marx's (1964) historical accounts of American antiurbanism.

been structurally transformed, rather than attenuated, in industrial bureaucratic social systems.

Community Saved

Many urban scholars have been dismayed by the Lost argument's emphasis on urban disorganization. In reaction, they have developed the Saved argument during the past 30 years, contending that neighborhood and kinship solidarities have continued to flourish in industrial bureaucratic social systems. The Saved argument asserts that such communal solidarities have persisted because of their continued efficacy in providing support and sociability, communal desires for informal social control, and ecological sorting into homogeneous residential and work areas (see Keller 1968; Suttles 1972). While granting that contemporary urban milieus also foster membership in more narrowly based multiple social networks, the Saved argument maintains that many of these networks tend to develop solidary features: single-stranded ties often broaden in scope as new aspects of the informal relationship develop (see Craven and Wellman 1973; Pickvance 1975), and densely knit, self-contained clusters of ties often emerge in initially sparse networks. Members of such networks are often important sources of assistance in mediating with formal bureaucratic structures and in coping with contingencies (e.g., Young and Wilmott 1957; Gans 1962, 1967; Liebow 1967; Stack 1974).[6]

Much of the Saved argument's case has rested on the sheer empirical demonstration of the continued vitality of those urban primary ties which had been pronounced Lost. Communal structures have been extensively documented in the Saved argument, in contrast with the Lost argument's analytic presentation of urbanites as aggregates of disconnected individuals. While some proponents of the Lost argument have alleged an association between communal disorganization and poverty (cf. Valentine's [1968] critique), those who have developed the Saved argument have found much evidence of solidary networks among poorer, traditional, or ethnic minorities seeking to maintain their resources against the claims of a centralizing state (cf. Tilly 1978). In the Saved argument, human beings are regarded as inherently gregarious, apt to organize communities under all circumstances. By the early nineteen sixties, the Saved argument had become the new orthodoxy, with the publication of such works as Gans's (1962) study of an "urban village," Greer's (1962) theoretical development of postwar survey research, and Jacobs's (1961) assertion of the vitality of dense, diverse central cities.

[6] There are clear similarities here to analyses of the importance of solidary ties in bureaucratic workplaces (e.g., Benyon 1973; Braverman 1974).

The problem of solidary ties.—Although the Lost argument's assertion of urban social disorganization has been rebutted, theoretically and empirically, this work has been accomplished by studies emphasizing the persistence of bounded communal solidarities. Such studies, while properly questioning the Lost argument's conclusions, have unfortunately not considered fully the Lost argument's useful starting point: that the contemporary division of labor may have strongly affected the structure of primary ties. Because Saved scholars have looked only for—and at—the persistence of communal solidarities in neighborhoods, in kinship systems, and on the job, it has been difficult to assess the position of solidary ties within overall social networks.[7] Weaker, more sparsely knit, more loosely bounded ties are all apt to be poorly represented in the Saved studies (see the discussion in Granovetter [1973]). While some Saved analyses have been quite concerned with external linkages, these linkages have been seen as radiating outward from a bounded communal base—often a small-scale territory or neighborhood (cf. Janowitz 1952; Greer 1962; Suttles 1972; Hunter 1975; Warren and Warren 1976; Warren 1978).

Thus the basic Community Question, dealing with the structure and use of primary ties, has been confounded in both the Lost and Saved arguments with questions about the persistence of solidary sentiments and territorial cohesiveness. But, whereas the Lost argument laments their demise, the Saved argument praises their persistence.

Community Liberated

The Liberated argument has developed out of the analytic juxtaposition of the Lost and Saved arguments. The Liberated argument affirms the prevalence and importance of primary ties but maintains that most ties are not now organized into densely knit, tightly bounded solidarities. The argument contends that: (*a*) the separation of residence, workplace, and kinship groups involves urbanites in multiple social networks with weak solidary attachments; (*b*) high rates of residential mobility weaken existing ties and retard the creation of strong new ones; (*c*) cheap, effective transportation and communication reduce the social costs of spatial distances, enabling the easy maintenance of dispersed primary ties; (*d*) the scale, density, and diversity of the city and the nation-state, in combination with widespread facilities for interaction, increase possibilities for access to loosely bounded, multiple social networks; and (*e*) the spatial dispersion of primary ties and the heterogeneity of the city make it less likely that those with whom an urbanite is linked will themselves be densely knit into solidary communities.

[7] Perhaps only Edward Banfield (1958) has gone out searching for solidary ties and not found any.

The Liberated argument has been systematically developed and tested only during the past 10 years.[8] Its take-off point has been that work of the Saved argument which has given analytic attention to urbanites' limited involvement in their local communities and to their external linkages beyond local boundaries (see also Merton 1957; Kasarda and Janowitz 1974; Taub et al. 1977). Taking this work further, the Liberated argument has abandoned the local area as the starting point for analyzing the Community Question and inquired directly into the structure of primary ties.

The Liberated argument contends that primary ties now tend to form sparsely knit, spatially dispersed, ramifying structures instead of being bound up within a single densely knit solidarity (see Kadushin 1966; Shulman 1972, 1976; Granovetter 1973; Laumann 1973; Breiger 1974; Shorter 1975; Fischer 1976; Walker 1977). While such ties may have fewer strands in the relationship than those in which kinship, residence, and work are combined, they are prevalent and important sources of sociability and support.

The argument suggests that primary ties are often dispersed among multiple, sparsely interconnected social networks. These networks, by their very nature, are not "institutionally complete" (Breton 1964), self-contained "urban villages." Their sparsely knit, ramifying structures provide a broad range of direct and indirect connections to the dispersed and differentiated resources of industrial bureaucratic social systems. Obtaining resources through such a sparsely knit network is not a matter of obligations due a member of a solidarity. Instead, it is a matter of the quality of the particular dyadic ties, the ease of maintaining contact, the ability of network members to provide indirect connections to additional resources, the extent to which additional members of a network can be mobilized to provide assistance, and the connectivity between networks (see Cohen 1969; Lee 1969; Bott 1971; Boissevain 1974; Granovetter 1974; Howard 1974; Walker 1974; Jacobson 1975; Fischer et al. 1977; Wireman 1978).

Answers and questions.—The Liberated argument has usefully freed the Community Question from its local roots. Yet a number of questions remain, because asserting that one should not set out initially to search for solidarities is quite different from asserting the nonexistence of such solidarities. First, to what extent do continuing kinship and local systems

[8] Some earlier scholars, who principally made the Lost argument, were also more optimistic at times about the consequences of this change in community structure. In their celebration of the potential for making choices among networks in the city, they prefigured the Liberated argument. For example, Georg Simmel contended that the urbanite, freed from a single encapsulating solidarity, had gained "freedom of movement . . . [and] a specific individuality to which the division of labor in the enlarged group gives both occasion and necessity . . ." ([1902–3] 1950, p. 417; see also [1908] 1971, p. 121). Robert Park's work (e.g., 1925*a*, [1925*b*] 1967) conveys a sense of excitement about the possibilities for individual action in the hurly-burly of the city.

structure primary ties? Second, are there no costs to maintaining ties over distance and no advantages to the quick physical accessibility afforded by proximity? Third, are there structural pressures toward the formation of solidarities, as friends of friends become friends of each other, as these increasingly dense clusters tend to interact more with each other and less across network boundaries, and as network members develop new strands in their relationships (see White 1965)? Fourth, are there circumstances—for example, lack of physical mobility or material resources, cultural differences—which can maintain dense, bounded solidarities? Fifth, is the maintenance of solidary sentiments dependent upon an unambiguous attachment to only one densely knit, tightly bounded communal structure?

Posing these questions is not to vitiate the Liberated argument, but to acknowledge that the formulation of the Community Question in network analytic terms has not only performed a useful critique of the Lost and Saved arguments but also provided us with a new structural perspective toward evaluating empirically some of their continuing concerns.

THE COMMUNITY QUESTION IN TORONTO

The foregoing analysis of the Community Question debate has developed concurrently with our research group's study of primary ties in Toronto. This research has been concerned with a number of issues which permeate the three arguments: To what extent are primary ties prevalent in industrial bureaucratic cities? To what extent is their composition based on kinship and neighborhood solidarities rather than on friendship? How homogeneous are urbanites' primary networks? How self-contained or ramified? How densely knit? What are the structural conditions associated with the availability of interpersonal assistance through these primary ties? The Lost, Saved, and Liberated arguments give quite different answers to these questions.

The Toronto research into such matters has been primarily survey based, supplemented by field work and focused interviews. The data discussed in this paper are derived from a 1968 random-sample survey of 845 adults (aged 18 and over) residing in the Toronto borough of East York. East York (1971 population = 104,645) is an upper-working-class/lower-middle-class, predominantly British-Canadian, inner suburb. Most residents live in small private houses or high-rise apartments; there are rarely more than two adults per household (see Gillies and Wellman 1968; Wellman 1976). East York has had the reputation of being one of the most solidary areas of Toronto. As such, it is a particularly interesting site at which to investigate the Community Question.

The survey asked respondents to provide detailed information about their six closest intimates ("the persons outside your home that you feel

closest to"), the ranked strength of closeness of their relationship with the respondent, their gender and socioeconomic status, the basis of their relationship (e.g., mother, neighbor), where they live, how often they are in contact (and by what means), and the kinds of assistance available in the relationships. Information was sought about the structure of these small egocentric intimate networks by inquiring into the respondents' reports of the interconnecting close ties among the sets of intimates named ($N =$ 3,930 intimates).[9]

Findings from this investigation of intimate ties will be presented in the next two sections. Despite the limitations of an analysis restricted to a quantitative case study of strong intimate ties, the data can help inform the Community Question debate. Research into the nature of primary ties is continuing in Toronto. Barry Leighton and I are now conducting in-depth reinterviews of a small subsample of the original respondents. A future monograph (Wellman, Shulman, Wayne, and Leighton, in preparation) will address such complementary matters as the nature of ties weaker than intimacy, the network dynamics of utilizing primary ties, longitudinal changes in primary networks (see also Crump 1977), and the relationship of solidary sentiments to network structures.

THE SOCIAL BASES OF INTIMACY

Relational Bases

Almost all (98%) East Yorkers report having at least one intimate tie; the majority (61%) report having five or more. Most have intimate ties with both kin and friends. For the sample taken as a whole, about half of all the intimate ties are with kin and about half are with unrelated individuals, predominantly "friends" who are not currently neighbors or co-workers (table 1).

The strongest intimate ties (in terms of the respondents' relative strength of closeness to these extrahousehold intimates) are usually with immediate kin (adult children, parents, and siblings), a traditional basis for solidary

[9] The questions to the respondents relevant for this paper were: "I'd like to ask you a few questions about the people outside your home that you feel closest to; these could be friends, neighbours or relatives. Please write in their initials, . . with the one you feel closest to on the first line, the next closest on the second line, and so on. Will you now tell me the relationship to you of each person you have written down. . . . Now, for the first person listed, . . Where does he/she live? How often do you *see* him/her? How often are you in touch by *phone* or *letter*? . . . Which of these do you rely on for help in everyday matters? Which of these do you rely on for help in an emergency? . . . [After six intimates:] I'd like to know which of the people . . . are close to one another. Tell me about the first one, please. Which of the others are close to that person? Which are close to Person 2? [etc.]" The data were originally collected in a study directed by Donald B. Coates, with Barry Wellman as co-director. See Coates et al. (1970, 1976). See also Wayne (1971).

TABLE 1

Strength of Intimate Relationship by Relationship to Respondent

Strength of Relationship (Ranked)	Relationship: (All Kin)	Child	Parent	Sibling	Other Relative	Friend	Neighbor	Co-worker	Total
1	525	106	153	167	99	226	32	28	811
	(27.1)	(46.7)	(45.0)	(28.2)	(12.7)	(15.3)	(13.1)	(13.0)	(20.9)
2	448	61	94	119	174	279	43	27	797
	(23.1)	(26.9)	(27.6)	(20.1)	(22.3)	(18.9)	(17.6)	(12.5)	(20.6)
3	352	30	38	128	156	298	55	36	741
	(18.2)	(13.2)	(11.2)	(21.6)	(20.0)	(20.2)	(22.4)	(16.7)	(19.1)
4	265	18	28	77	142	280	38	43	626
	(13.7)	(7.9)	(8.2)	(13.0)	(18.2)	(19.0)	(15.5)	(19.9)	(16.2)
5	193	8	20	52	113	227	42	46	508
	(10.0)	(3.5)	(5.9)	(8.8)	(14.5)	(15.4)	(17.1)	(21.3)	(13.1)
6	155	4	7	49	95	166	35	36	392
	(8.0)	(1.8)	(2.1)	(8.3)	(12.2)	(11.2)	(14.3)	(16.7)	(10.1)
N	1,938	227	340	592	779	1,476	245	216	3,875
% of total	50.0	5.9	8.8	15.3	20.1	38.1	6.3	5.6	100.0
Mean rank	2.8	2.0	2.1	2.8	3.4	3.3	3.5	3.7	..

NOTE.—$\chi^2 = 441.3$, $P < .001$; gamma = .27; nos. in parentheses are percentages; missing data = 55.

ties. Furthermore, when neighbors and co-workers are considered as intimates at all, the ties with them are likely to be comparatively weak (table 1).

Most East Yorkers specialize in one type of intimate relationship, kin or friend, but they also maintain one or two other types of intimate ties. A sizable minority are "superspecialists": 19% name only kin and 18% name only nonkin. Kin and nonkin intimates tend to be in different clusters of their intimate networks and not to have intimate ties with each other. All of an East Yorker's intimates, though, are indirectly tied to each other through the respondent; many may also have nonintimate direct connections with one another.

The multiple bases of the intimate ties (kinship, friendship, etc.) and the lack of direct connections between the relationally different intimates are in accord with the Liberated argument (see Laumann 1973; Verbrugge 1977; Fischer et al. 1977). Yet multiplicity does not mean equality. Most East Yorkers feel closer to kin than to unrelated intimates, and the greater number of their intimate ties tend to be bound up in one type of relationship.

The prevalence and importance of kinship ties is congruent with the Saved argument (e.g., Litwak 1960; Adams 1968; Klatzky 1971; Gordon 1977). However, in treating kinship systems as separate analytic entities, such Saved arguments may have underplayed the multiple bases of contemporary urban intimate networks. Our data suggest a synthesis of the Liberated and the Saved arguments: the variety of intimate ties potentially provides access to a more diverse array of resources, while heavy involvement with kin retains connections to a somewhat solidary system.

Spatial Expanse

The distribution of intimates' residences reveals that these strong primary ties of East Yorkers are situated in a broad field of interaction in Metropolitan Toronto and beyond. The great majority of East Yorkers' intimates live within Metropolitan Toronto, but only a small minority (13%) live in the same neighborhoods as their respondents (table 2). The metropolitan area thus bounds the effective field of interaction more than does the neighborhood. However, one-quarter of the intimates live outside Metropolitan Toronto, some as far away as Vancouver and New Delhi.

The distances at which intimate links are apt to be maintained vary markedly with the relational basis of the tie. Distant ties are much more likely to be with kin than with friends: 34% of intimate kin live outside Metropolitan Toronto, more than twice the percentage of unrelated intimates (table 2). Furthermore, ties with kin are the more actively main-

TABLE 2

RESIDENCE OF INTIMATES BY RELATIONSHIP TO RESPONDENT

	RELATIONSHIP							
RESIDENCE	Child	Parent	Sibling	Other Relative	Friend	Neighbor	Co-worker	TOTAL
Same neighborhood	9 (4 0)	23 (6 8)	25 (4 2)	54 (6.9)	194 (13 1)	182 (74 3)	18 (8 3)	505 (13 0)
Elsewhere in East York	23 (10 1)	35 (10 3)	63 (10 6)	85 (10 9)	211 (14 3)	38 (15.5)	28 (13 0)	483 (12.5)
City of Toronto	26 (11.5)	94 (27 6)	130 (22 0)	176 (22 6)	441 (29 9)	10 (4.1)	83 (38 4)	960 (24 8)
Elsewhere in Metro Toronto	108 (47.6)	50 (14 7)	147 (24.8)	227 (29.1)	359 (24 3)	10 (4.1)	71 (32 9)	972 (25 1)
Outside Metro Toronto	61 (26 9)	138 (40 6)	227 (38 3)	237 (30.4)	271 (18.4)	5 (2.0)	16 (7 4)	955 (24 6)
N	227	340	592	779	1476	245	216	3875
% of total	5 9	8 8	15 3	20 1	38.1	6 3	5.6	100 0

NOTE.—$\chi^2 = 1203.0$, $P < .001$; gamma = −.32 (relationship ordered by mean closeness rank for kin and nonkin [see table 1]); nos in parentheses are percentages.

tained distant intimate ties, with a much higher frequency of in-person and telephone interaction.

The wide spatial expanse of intimate networks is facilitated by the telephone. Indeed, telephone contact between intimates is usually more frequent than is in-person contact (table 3). The two modes of communication are generally complementary and not substitutive; it is quite rare for there to be a good deal of telephone conversation between intimates without there also being frequent in-person meetings.[10]

Perhaps the greater bandwidth of communication available through in-person meetings provides necessary information to reaffirm, reinforce, and readjust relationships maintained routinely by telephone. In no instance is an intimate tie sustained solely through telecommunications.

Distant ties.—Contemporary transportation and communication facilities have lessened, but not eliminated, the constraints of distance on maintaining intimate contact. Intimates who live far from East York tend to have a different relationship, having much less frequent telephone and

TABLE 3

RELIANCE ON DIFFERENT MODES OF CONTACT, CONTROLLED BY INTIMATES' RESIDENTIAL LOCATION (%)

	Residential Location				
Mode and Frequency of Contact	Same Neighborhood	Elsewhere in East York	City of Toronto	Elsewhere in Metro Toronto	Outside Metro Toronto
In person weekly or more often; telephone weekly or more often	51.9	53.4	43.0	38.8	6.8
In person weekly or more often; telephone twice per month or less	31.5	15.5	13.1	11.6	4.8
In person twice per month or less; telephone weekly or more often	4.8	9.5	20.9	20.8	21.8
In person twice per month or less; telephone twice per month or less	11.9	21.5	23.0	28.8	66.7
Total	100.1	99.9	100.0	100.0	100.1
N	505	483	975	984	947
χ^2 ($P <$ .001)	31.0	95.2	85.9	123.8	53.5
Conditional gamma (in person by telephone for each residential location)	.61	.77	.56	.64	.63

NOTE.—Zero-order gamma (in person by telephone, uncontrolled by residential location) = .67; partial gamma (in person by telephone, controlled by residential location) = .62

[10] One straightforward exception is that intimates living on the same block rarely use the telephone for contact.

in-person contact.[11] Their infrequent contact ratifies the tie, and a potential is retained for more intensive use when needed. The minority of those distant intimates who do interact frequently tend to maintain contact by telephone (table 3).

Clearly, many of the long-distance intimate ties are rather dormant in their actual functioning, maintained through infrequent contact and structural embeddedness (see also F. Katz 1966; P. Katz 1974). However, the very existence of these semidormant ties may usefully link East Yorkers to other connections. Furthermore, these are intimate ties and not just distant links to kin and friends. There is the memory of past interaction and the anticipation of future use. When necessary, the costs of distance can be overridden by an emergency, as when a respondent flew 2,100 miles to nurse a sick mother in Calgary although Sunday telephone calls had sufficed for the previous 10 years (see also Boswell's [1969] Zambian example).

Local ties.—The great majority of East Yorkers' intimate networks are not organized into local solidarities. Few have more than one intimate who resides in their own neighborhood.

Yet East York's pride in its local community ties is not without foundation. Although the borough contains less than 5% of Metropolitan Toronto's population, fully one-quarter of the respondents' intimate ties are to other East Yorkers, and the percentage is even higher for ties to intimates who are not kin (see table 2). Furthermore, many now-distant ties had local origins (Shulman 1972).

Most East Yorkers also have useful ties with neighbors, although these rarely reach the strength of intimacy. On the average, they talk with five neighbors regularly and visit in the homes of three (Gates, Stevens, and Wellman 1973). Such local ties are used for easy sociability and assistance when quick physical accessibility is an important consideration.

The data on the spatial expanse of intimacy provide support for a synthesis of the Liberated and the Saved arguments. East York is neither a Gansian "urban village" nor a "community without propinquity" (Webber 1963). While local ties are real and important, their importance comes from their being only a component of a diverse array of relationships. Intimate ties are organized into local solidarities even less often than they are into solidary kinship systems. Indeed, the car, the telephone, and the airplane help maintain many kinship ties. Yet space is still a constraint; there are distances for each tie at which the cost of keeping in contact becomes too great for it to remain viable.

[11] Indeed, the nature of the relationships may affect the spatial expanse of the tie, as when an aging mother decides to rent an apartment near her daughter.

Network Structure

Density.—The mean density of East Yorkers' intimate networks is 33%; that is, one-third of all possible intimate ties between respondents' intimates are actually reported to exist. Only one-fifth of the networks have a density greater than 50% (table 4), although many intimates are more weakly connected to each other (cf. Granovetter 1973). Thus, the great majority of respondents are not encapsulated within the bounds of one solidary group, but are linked through their intimates to multiple, not strongly connected, social networks. The prevalent sparse density supports the Liberated argument.[12]

There are significant clusters of density within networks, though. Kinship systems often foster close ties among members, and those intimate networks which are predominantly composed of kin tend to be more densely knit than the others (see table 4). Kin members of intimate networks also tend to form densely knit clusters within the rather sparse overall networks. Intimate friends, in contrast, tend either to be unconnected to other intimates or to be linked dyadically to them.

Reciprocity.—Shulman's associated study (1972, 1976) interviewed 198 of the intimates named by a subsample ($N = 71$) of our respondents and

TABLE 4

KIN AND FRIENDS IN INTIMATE NETWORKS BY DENSITY

Density of Networks (Grouped) (%)	N	%	Cumulative %	% Kin in Such Networks	% Friends in Such Networks
0–25	388	47.1	47.1	36.4[a]	53.2[a]
26–50	261	31.7	78.8	56.9	35.9
51–75	65	7.9	86.7	56.9	37.0
76–100	110	13.3	100.0	73.7	20.1
Total	824	.	.	49.5	42.1
F(3,820)	.	.		48.6*	36.6*

[a] Rows do not add to 100% because co-worker and neighbor intimates are not included.
* $P < .01$.

[12] This is called "sparse density" (or "sparsely knit") because less than a majority of all possible interconnections actually appear. However, Harrison White has pointed out that, without standards for comparison, we have no firm theoretical or empirical grounds for expecting higher density, especially when studying strong ties of intimacy (personal communication). Jack Wayne, using the same procedure as employed here, found the density of reported ties between intimates in an inland Tanzanian area (Kigoma) to be 76% (personal communication). Ties between respondents and intimates have been excluded in the density calculations, as such ties exist by definition. Links were calculated symmetrically: if a respondent reported intimate no. 1 to be close to intimate no. 2, it was also assumed that he or she perceived intimate no. 2 to be close to intimate no. 1.

asked them, in turn, who their intimates were. Overall, only 36% of the surveyed intimates reciprocally named East Yorkers as their intimates. The closest intimates (those ranked first by mutual respondents) were markedly more likely to see each other as mutual intimates. Others acknowledged return ties to the East Yorkers but weaker ones than intimacy. They have intimate relationships but different ones from those of the East Yorkers. These ramifying, nonreciprocating ties are in keeping with the Liberated argument and argue against the Saved argument's notion of tightly bounded, mutually oriented solidarities.[13]

Ramifications.—Taken together, the variety of types of intimate ties, the sparse network density, and the often-unreciprocated character of intimate bonds strongly suggest a ramified, loosely bounded web of primary ties, rather than an aggregation of densely knit, tightly bounded solidary communities. Only a minority of an East Yorker's intimates reciprocate intimacy, and only a minority of intimates are reported to be intimate with each other. The overall structure of intimate relationships is in accord with the Liberated argument.

Yet the data also indicate some basis for the closer structural integration suggested by the Saved argument. There are often dense clusters within more sparsely knit networks. Furthermore, many of those who are not intimately connected with each other are linked together in other important ways: as friends, acquaintances, neighbors, co-workers, and nonintimate kin. Using such less restrictive criteria, there is much structural connectivity.

THE AVAILABILITY OF HELP FROM INTIMATES

If East Yorkers are to avoid the direct dependence on formal bureaucratic resources seen by the Lost argument, they must be able to obtain assistance through their primary ties. Although such assistance might come through many ties, it is reasonable to expect that much reliance would be placed on help from intimates, the people outside the household to whom they feel closest.[14]

13 Shulman's findings (1972, 1976) indicate that the symmetry assumption may well overstate the density of the networks when only intimate ties are considered, although we wonder if the respondents would have perceived the asymmetry present in the ties between their intimates. This lack of reciprocity gives a structural basis for expecting wide disparities in the extent to which an urbanite is chosen as an intimate (see Rapoport and Horvath's [1961] study of a biased friendship network). This, in turn, indicates the structural prevalence of "brokerage" nodes, whose heavily chosen incumbents link together a number of social networks.

14 Our in-depth interviews indicate that respondents often perceive their intimate connections as a type of general utility. While they know that they might need help from intimates at some time, and maintain their ties in part for that purpose, often they do not have any precise idea of what contingencies will in fact develop. The treatment

The structural situation of East Yorkers, linked to intimates by means of sparse networks rather than solidary groups, is reflected in the nature of the help reportedly available. In support of the Liberated argument, we find that the great majority of East Yorkers (81%) report that help in emergency situations is available to them from somewhere in their intimate network. A smaller majority (60%) report help to be available through their intimate networks in dealing with everyday matters; such routine help is often available as part of less intensive relationships (e.g., with acquaintances, neighbors, co-workers), and there is less use of intimate relations for it.

While assistance in emergencies is available from the great majority of intimate networks, it is not available from the majority of intimate relationships. Only a minority of intimates, 30%, help in emergencies, and only 22% help in dealing with everyday matters. Thus, East Yorkers can almost always count on help from at least one of their intimates, but they cannot count on such help from most of them.

To some extent these data are consistent with the Lost argument's concerns about the attenuation of supportive primary ties. However, the data support more fully a differentiated conceptualization of intimacy, consistent with the Liberated argument's analysis of the division of labor in primary networks. Intimacy (or closeness) is not a unidimensional construct. "Helping" is a defining attribute of a minority of intimate relationships, while others may be based on sociability, structural or normative obligation, or propinquity (see Leyton 1974). The remainder of this section examines the effect of relational and structural factors in these differentiated networks on the likelihood that a respondent regards an intimate as a provider of assistance. Two path diagrams summarize the interrelationships; table 5 presents the correlation matrix for both diagrams.[15]

of help as a generalized resource is a conservative estimate of its availability from intimates. East Yorkers may count on help from some other intimates for specific contingencies, defined by the relationship and the resources available, while not thinking of these intimates as being generally helpful. It is the generalized role relationship of "helper," clearly evident in our in-depth interviews and field work, that is of interest here Such membership in general-purpose helping relationships challenges a market model of assistance, in which a seeker rationally determines a need, scans all available sources, and calls upon them in ranked order of probable utility. Not only is the provision of help determined by networks, but so may be the perception and utilization of available channels. Indeed, the very provision of help may precede—and define—the putative seeker's desire to enter into a help-receiving relationship.

15 Earlier analyses (e.g., Wellman et al. 1971) indicate that there are no appreciable direct associations between such social categorical variables as age and SES and the availability of help They thus were omitted from the analysis for this paper. See Wellman (1977) for cross-tabulations presenting more detailed information about the relationships between the network variables and the availability of help.

TABLE 5

ZERO-ORDER CORRELATION MATRIX FOR PATH DIAGRAMS

	Parent/ Child	Lives inside Metro Toronto	Co-worker	Centrality	In-Person Contact	Closeness Strength	Telephone Contact	Emergency Assistance
Lives inside Metro Toronto.....	−.100							
Co-worker (fig. 2 only)..	− 102	101						
Centrality........	236	− 079	− 092					
In-person contact	− 028	.346	437	− 012				
Closeness strength........	.267	− 011	−.098	044	035			
Telephone contact.........	205	282	.008	098	.337	.212		
Emergency assistance from intimate (fig. 1)........	.180	.105	− 021	054	.147	280	197	
Everyday assistance from intimate (fig. 2)	.112	134	.093	065	202	.221	.248	.444

Kinship

The most antecedent variable in the path models traces the continuing effect of kinship ties. The role of the extended family as a special provider of assistance is confined among East Yorkers to intimate parents and (adult) children. Other intimate kin, such as siblings, grandparents, and aunts, are only as likely as friends to provide assistance.

Parent-child support is more marked in emergencies: 50% of parent-child ties have an emergency-assistance component, as compared with 26% of other intimate ties. Parents and children are more apt than other intimates to be called upon for help in emergencies, regardless of where they live (they tend to live at greater distances from respondents than other intimates) and how frequently they are in face-to-face contact (fig. 1). In addition to the direct effect, intimate parents and children are also more likely than other intimates to provide help, because they tend to have closer ties with the East Yorkers and to be in more frequent contact with them.

While parent-child intimates are also significantly more likely (34%) than all others (19%) except co-workers (37%) to help out with everyday affairs (Wellman 1977), there is no direct effect on kinship in this case (fig. 2). There are indirect effects, however, due to the stronger bonds and more frequent telephone contact that parent-child intimates have.

The kinship data partially support both the Saved and the Liberated models. Kinship remains a significant basis for providing help, both directly and because it encourages closer bonds and more frequent telephone contact. Yet the particularly helpful intimates are parents and children and not a large solidary network of extended kin relations.

Propinquity

Our earlier analyses (Wellman 1977) showed the availability of assistance to be not significantly associated with intimates' neighborhood residence, in contradistinction to the Lost and Saved arguments' emphases on local solidarities. Accordingly, the local residence variable has been omitted from the final path analyses.

Proximity appears to be more important on the job than in the neighborhood for the availability of help from intimates. Co-workers' frequent face-to-face contacts make them a significant source of everyday assistance for East Yorkers, despite the comparative weakness of their intimate bonds.

The residential distinction that does make a difference in the availability of help is that of living inside Metropolitan Toronto's boundaries; that is, being a local call or a short drive away. This has a slight direct positive effect on the availability of help and appreciably increases the fre-

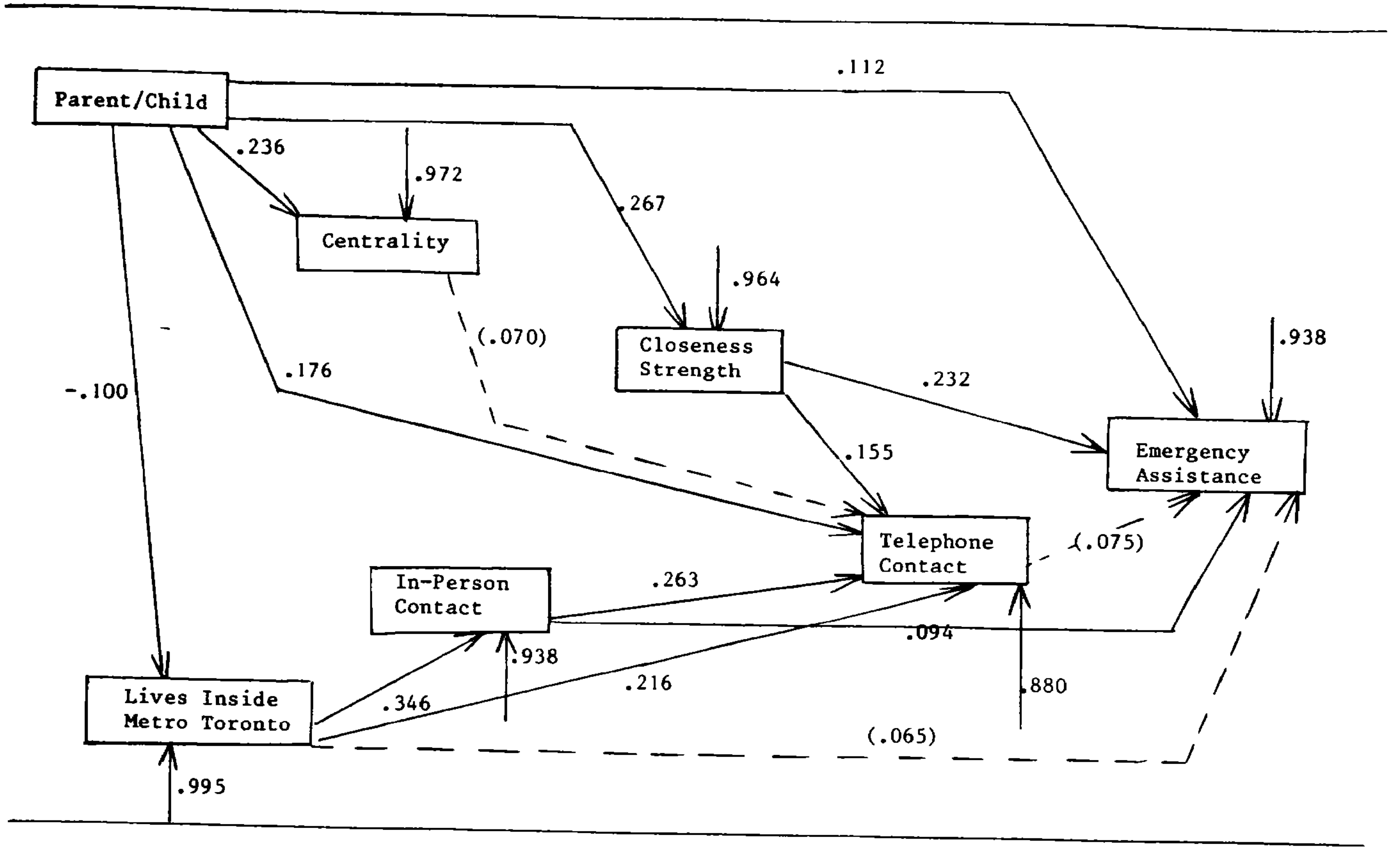


Fig. 1.—Path model: emergency assistance from intimates. Dashed lines with correlations in parentheses represent nonsignificant paths necessary to reproduce original correlation matrix to .05.

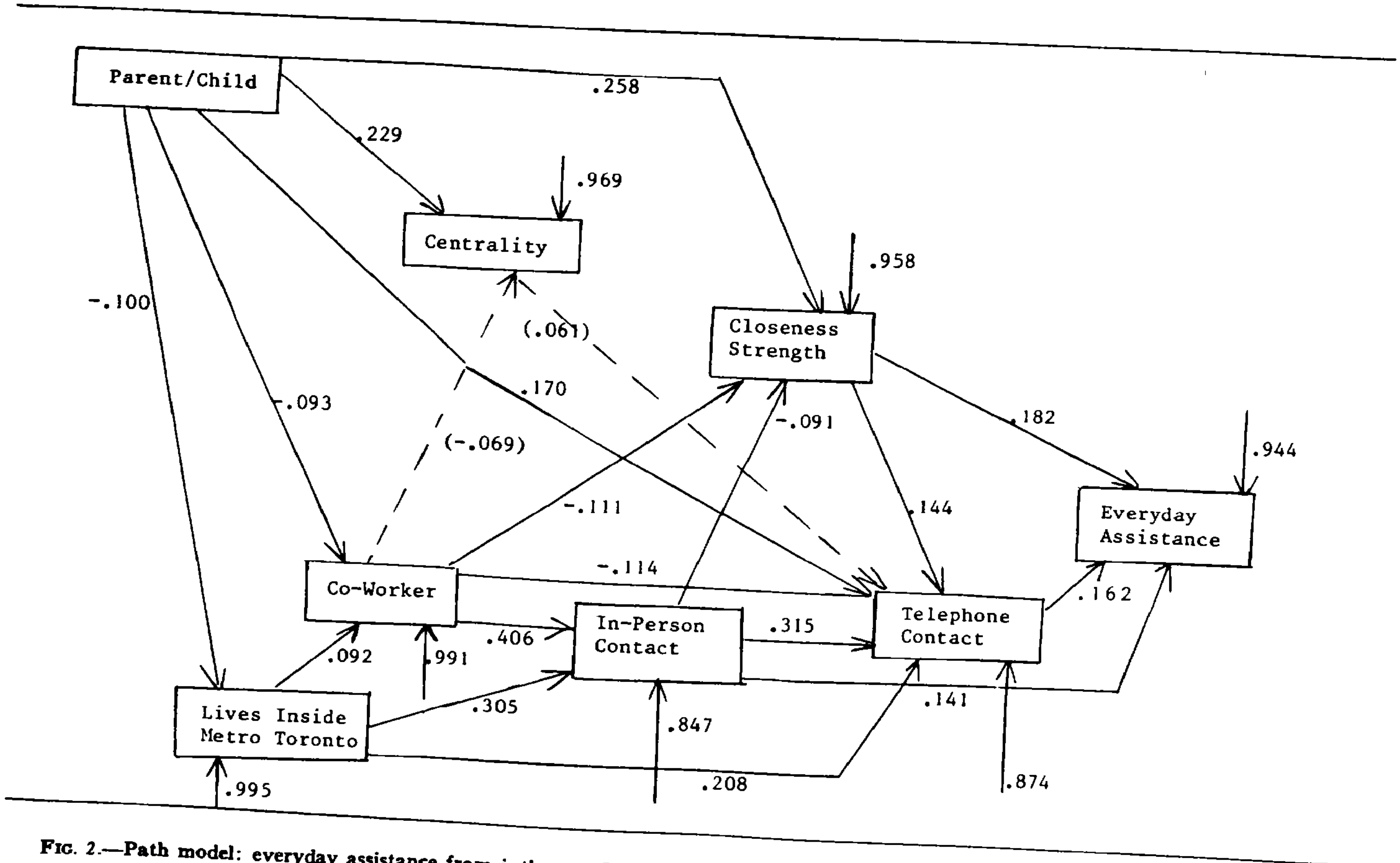


FIG. 2.—**Path model: everyday assistance from intimates. Dashed lines with correlations in parentheses represent nonsignificant paths necessary to reproduce original correlation matrix to .05.**

quency of contact between intimates. (There may be reverse effects operating as well, with intimates choosing to live in Metropolitan Toronto so that they may continue to be available to help their East York respondents.) The data support a somewhat revised version of the Liberated argument: to an appreciable extent, the spatial range of assistance relationships has not disappeared, but has expanded to encompass the entire metropolitan area.

Centrality and Density

One purely structural variable, an intimate's centrality in a respondent's network (measured as the number of intimate ties that an intimate has with any of the respondent's other intimates), slightly affects the frequency of telephone contact and, hence, the provision of assistance. In general, more structurally central intimates are more likely to provide help. Indeed, their ability to provide help may have made them central.

A structurally central person's potential ability to mobilize help is not related to the solidary nature of the network: no significant paths between the density of a respondent's network of intimates and the availability of assistance from an intimate have been found in our analyses. Hence density has been deleted from the final path models. The absence of significant density effects and the weak effects of centrality also argue that the helpfulness of parents and children is independent of the potential solidarity of their kinship networks. It is a component of dyadic parent-child relations. In sum, the centrality and density data support the Liberated argument better than the Saved argument.

Frequency of Contact

The more frequently intimates are in contact, especially in person, the more apt are they to provide assistance in their relationships.[16] Frequent contact is particularly associated with the more mundane provision of everyday assistance, when ready accessibility is more likely to be a mobilizing factor.

Closeness

The closer (stronger) the intimate relationship (as measured by the respondent's ordinal ranking of the intimates), the more the perceived

[16] For the path analyses, the original categorically recorded frequency-of-contact data were transformed into estimated-days-per-year equivalents E g , "about once a week" was transformed into "52." This transformation makes the simplifying assumption of equal time spent per contact.

availability of help becomes a salient defining component of that tie. Closeness is apparently the single most important defining characteristic of helpful intimate relationships; it is the strongest direct predictor in the path models.[17] For example, 56% of the first closest ranked intimates are relied on in emergencies by East Yorkers, while only 16% of the sixth closest intimates are. Closeness also has appreciable paths to the next most powerful predictors, the frequency of contact variables. Furthermore, all other significant variables predict to it, directly or indirectly.

The data indicate that the availability of help to East Yorkers from intimates is a process more fully in accord with the Liberated argument than with the Saved argument. The full path diagrams show two social processes, both more closely associated with the nature of two-person intimate bonds than with the structure of overall intimate networks. On the one hand, a comparatively strong "interactional" set of paths go from the spatially propinquitous facilitation of interaction (through living in the same metropolitan area or being a co-worker) to the frequency of interaction to the availability of assistance. On the other hand, another set of "familial" paths go from parent-child ties to the strength of closeness of intimate ties to the availability of assistance.

The availability of the parent-child tie for assistance is not associated with the tie being embedded in strong, supportive kinship relationships. Furthermore, structural variables, such as centrality and density, are poorly related to the availabiilty of assistance. The availability of assistance thus is more closely associated with the character of the two-person bond than it is with the potentially solidary character of the overall network.

Although the data document network effects on the availability of help from intimates, the amounts of explained variance in the path models are not large. I am reluctant to relinquish most of the unexplained variance to unspecified, residual "psychological factors." Some of the unexplained variance is probably due to the crude way in which the variables have been defined and measured. Furthermore, the way is surely open to the delineation of additional structural and categorical variables that can affect the interpersonal provision of scarce resources.

COMMUNITY IN EAST YORK: LOST, SAVED, OR LIBERATED?

Community Lost?—The prevalence of strong intimate ties in East York calls into question the basic contention of the Lost argument (see summary table 6). If kin and neighbors have been lost as intimates, they apparently

[17] We are concentrating here on predicting to the reported availability of help. But it is also quite likely that, reciprocally, being perceived as helpful may engender stronger perceptions of closeness among intimates.

TABLE 6

THE COMMUNITY QUESTION: LOST, SAVED, AND LIBERATED ARGUMENTS COMPARED WITH EAST YORK FINDINGS

Argument	Community Lost	Community Saved	Community Liberated	East York Findings (Main Tendencies)
Basis of intimacy:				
Availability	Rare	Abundant	Abundant	5+ intimates
Relational	Formal role	Kin, neighborhood	Friendship, work	Kin, friendship
Spatial	Local	Local	Metropolitan, national	Metropolitan
Mode of contact	In person	In person	In person, telephone	Telephone, in person
Communal structure:				
Density	Sparse	Dense	Sparse	Sparse
Reciprocity	No	Yes	Uneven	Uneven
Boundedness	Ramified	Tight	Ramified	Ramified
Basis of assistance:				
Prevalence	Minimal	Abundant	Moderate	Moderate
Relational source	Formal ties	Kin, neighborhood	Friendship, work	Parent/child, work
Residential basis	Local*	Local	Metropolitan, national	Metropolitan
Density	Dense*	Dense	Sparse	N S.
Structural source	Secondary	Solidary group	Network ties	Network ties

* To the extent to which primary ties exist.

have been replaced by friends and co-workers. Yet East Yorkers report that they can count on only a minority of their intimates for help. Communal networks of mutually supportive intimate relationships do not appear. If highly supportive communities ever did exist for East Yorkers, intimate ties now occur only as much more differentiated networks. However, in these networks, many intimate ties contain support as an important strand in the relationship, and help from intimates is available to almost all East Yorkers.

Community Saved?—The data support some aspects of the Saved argument, albeit greatly affected by the contemporary context (see summary table 6). Parent-child ties play a special role in the overall intimate networks. They tend to be socially closer than other intimate ties, even at greater physical distances. Parents and children are more apt to provide help in mundane matters as well as in crises. Other intimate kin, however, can be counted on no more than can intimate friends. Clearly, the important kinship obligations that most intimate parents and children maintain operate as dyadic relationships, as the data also indicate that the density of the network is not a factor in the mobilization of assistance. There are few large solidary networks of helpful kin.

Residential propinquity still facilitates the provision of assistance, but

the local area is now metropolitan and not the neighborhood. This implies that it is the physical availability of aid—by automobile, public transit, and telephone—which is operative and not the activity of neighborhood solidarities. Kinship and metropolitan residence both act to encourage frequent contact. Those in contact more often are more likely to feel closer and to provide assistance when needed.

Neighborly relations are prevalent and, for many East Yorkers, important. Considered separately, they validate the opinion East Yorkers have of themselves as being heavily involved in local community interactions. Yet the data indicate that such neighborhood ties are usually just one component of a more diverse set of relationships and that they rarely comprise the more intense intimate relationships. It may be that in other populations, less mobile or less preoccupied with controlling internal resources, a larger proportion of the relationships will be tightly bound in solidary groups (see Wolf 1966; Wellman and Leighton 1979). Clearly, more work needs to be done on the dynamics of establishing, maintaining, using, transforming, and losing primary ties.

Community Liberated?—Our findings most fully support the Liberated argument that East Yorkers tend to organize their intimate relationships as differentiated networks and not as solidarities. There is much differentiation in the nature and use of intimate ties. There are links to a variety of people with different structural positions, often living in quite different residential areas (or interacting at work), and maintaining contact both by telephone and in person, at a wide range of time intervals.

The availability of assistance is affected by the quality of the relationships and not by the extent of structural solidarity. Not all intimate ties are used similarly, even those which are densely knit. Some intimates can be counted on to provide assistance in dealing with everyday matters; a good many more, but certainly not all, give assistance when emergencies arise. Other intimates interact with East Yorkers on different bases, such as kinship obligations, sociability, or job comradeship.

The many components of intimate relationships are not very neatly associated; the "role frame" of intimacy includes many complexly packaged bundles of relationships (see Nadel 1957). Some intimate friends are seen only socially; some provide help routinely. Some intimate kin can be counted on in any emergency; others cannot. Some intimate kin are seen daily; some intimate friends, yearly; and so on. This lack of neat coincidence among the qualitative aspects of the relationships makes for more differentiated linkages (see Mitchell 1969; Litwak and Szelenyi 1969; Gordon 1976). Different contingencies, social situations, and times of the day, week, or year bring East Yorkers into juxtaposition with a variety of connections.

The data analyzed here may help to resolve the discrepancy between the

Saved argument's extensive documentation of solidarities and the Liberated argument's portrayal of differentiated networks. The resolution may be a matter of analytic scope. If one focuses on kinship systems or neighborly relations, one is apt to find densely knit, tightly bounded networks. Looked at in fine-grain isolation, these networks appear as solidarities, which may well serve to give urbanites a sense of attachment in the social system. But if one broadens one's field of view to include all those with whom an urbanite is in touch, then the apparent solidarities may be seen as clusters in rather sparsely knit, loosely bounded networks.

COMMUNITY: SOLIDARITY OR NETWORK?

Intimate networks are just one of a number of often quite distinct personal networks. Frequently, weaker ties, such as neighboring and coworking, have limits on the claims that can be made on them. But they also tend to provide indirect access to a greater diversity of resources than do stronger, more socially homogeneous ties (Granovetter 1973).

All persons with whom one is directly connected are indirectly linked to each other through oneself. Each individual is a member of the unique personal networks of all of the people with whom he or she is linked, and membership in these networks serves to connect a number of social circles (see Craven and Wellman 1973). Thus, complex networks of chains and clusters are ultimately connected via a common network node. Social solidarity, analyzed from this perspective, may be the outgrowth of the coordination of activities through network processes rather than of the sharing of sentiments through common socialization.

While making for low communal solidarity, a variety of ties and uneven network density provide structural bases for dealing with contingencies. Densely knit network clusters can provide the basis for cooperative activities. Ramifying networks and asymmetrically reciprocated linkages can facilitate access to other social circles.

The concatenation of networks helps to organize social systems. Considered from the standpoint of the system rather than from that of the individual, it is the compounding of links and networks at many levels which allocates resources and juxtaposes alliances of similar interests. Not only individuals, but also clusters and collectivities, are linked through network ties (cf. Granovetter 1976; Rytina 1977; Laumann, Galaskiewicz, and Marsden 1978). A network of networks connects individuals, clusters, and collectivities in complex ways.

Despite all this connectivity, our data also suggest why so many urbanites believe in the Lost argument, even when they themselves are well connected. Rather than an unambiguous membership in a single, almost concrete, solidary community, East Yorkers' lives are now divided among

multiple networks. The sparseness of interconnections among those networks means that no one solidarity can readily make or enforce general claims on a member. While this may be somewhat liberating in providing structural room to maneuver, it may also create a disorientating loss of identity, as it is no longer as clear or simple to which group (among many) one belongs. Although urbanites have not lost their communal access to people and resources—and, indeed, may have increased their reach—for those who seek solidarity in tidy, simple hierarchical group structures, there may now be a lost sense of community.

Yet membership in spatially and socially ramified networks is a useful way for urbanites to have access to diverse and differentiated resources not available through solidary auspices. Their ties are not encapsulated in "decoupled" little worlds (White 1966) but are strands in the larger metropolitan web. The ties provide the basis for network members to utilize the connections with others that their alters have. This suggests that Liberated networks may be more than just a passive rearrangement of primary ties in response to the pressures of large-scale social system changes. Instead, they can well be active attempts by contemporary urbanites to gain access to and to control system resources, given differentiated social systemic divisions of labor.[18]

REFERENCES

Abu-Lughod, Janet 1969. *The City is Dead—Long Live the City* Berkeley: Center for Planning and Development Research, University of California.

Adams, Bert N. 1968. *Kinship in an Urban Setting*. Chicago: Markham

Banfield, Edward. 1958. *The Moral Basis of a Backward Society*. New York: Free Press.

Barnes, J. A 1972. *Social Networks*. Reading, Mass.· Addison-Wesley.

Bender, Thomas. 1978. *Community and Social Change in America*. New Brunswick, N.J.: Rutgers University Press.

Benyon, Huw. 1973. *Working for Ford*. London: Allan Lane.

Boissevain, Jeremy F. 1974. *Friends of Friends*. Oxford: Basil Blackwell.

[18] The discussion in this paper has been limited to the contemporary industrial bureaucratic situation, and no attempt has been made to address complex cross-cultural and longitudinal issues. Yet a developmental perspective that sees industrialization, bureaucratization, and urbanization as inevitably fostering communal transitions from Saved or Lost structures is quite questionable. It is entirely possible that Liberated networks exist in nonindustrial milieus under appropriate structural conditions of resource access and control. On the one hand, recent studies clearly indicate that differentiated long-distance ties are maintained in the Third World without the prevalence of telephones or private automobiles as long as the ties are structurally embedded in kinship systems or common local origins. A number of mechanisms are used to maintain contact, such as trips by buses and trucks, messages sent with covillagers or travelers, and hired letter writers and readers (e g, Cohen 1969; Jacobson 1973; Howard 1974; Mayer and Mayer 1974; Ross and Weisner 1977; Weisner 1973). On the other hand, historians have been reporting nonsolidary aspects of many preindustrial Western European and North American communities, with status-group cleavages, appreciable spatial mobility, and complex households having heterogeneous external ties (e.g., Laslett 1971; Scott and Tilly 1975; Shorter 1975; Tilly 1975; Bender 1978).

Boswell, David. 1969. "Personal Crises and the Mobilization of the Social Network." Pp. 245–96 in *Social Networks in Urban Situations,* edited by J. Clyde Mitchell. Manchester: University of Manchester Press.

Bott, Elizabeth. 1971. *Family and Social Network.* 2d ed. London: Tavistock.

Braverman, Harry. 1974. *Labor and Monopoly Capital.* New York: Monthly Review Press.

Breiger, Ronald. 1974. "The Duality of Persons and Groups." *Social Forces* 53 (December): 181–89.

Breton, Raymond. 1964 "Institutional Completeness of Ethnic Communities and the Personal Relations of Immigrants." *American Journal of Sociology* 70 (September): 193–205.

Castells, Manuel. 1976. "Is There an Urban Sociology?" Pp. 33–59 in *Urban Sociology: Critical Essays,* edited by C. G Pickvance. London: Methuen.

Coates, Donald B., and associates. 1970. "Yorklea Social Environment Survey Research Report." Mimeographed. Toronto: Clarke Institute of Psychiatry, Community Studies Section.

Coates, Donald B., S. Moyer, L. Kendall, and M. G. Howatt. 1976. "Life-Event Changes and Mental Health." Pp. 225–49 in *Stress and Anxiety,* edited by C. D. Spielberger and Irwin G. Sarason. Vol. 3. New York: Halsted Press.

Cohen, Abner. 1969. *Custom and Politics in Urban Africa.* Berkeley: University of California Press.

Craven, Paul, and Barry Wellman. 1973. "The Network City." *Sociological Inquiry* 43 (December): 57–88.

Crump, Barry. 1977. "The Portability of Urban Ties." Paper presented at the annual meeting of the American Sociological Association, Chicago, September.

Death Wish. 1974. Directed and coproduced by Michael Winner. Written by Wendell Mayes. Produced by Hal Landers and Bobby Roberts. A Dino De Laurentis production From the novel by Brian Garfield.

Emerson, Richard. 1962. "Power-Dependence Relations." *American Sociological Review* 27 (February): 31–41.

Feagin, Joe. 1973. "Community Disorganization" *Sociological Inquiry* 43 (Winter): 123–46.

Fischer, Claude S. 1976. *The Urban Experience.* New York: Harcourt Brace Jovanovich.

———. 1977. "The Contexts of Personal Relations: An Exploratory Network Analysis." Working Paper no. 281. Berkeley: Institute of Urban and Regional Development, University of California.

Fischer, Claude S., Robert Max Jackson, C. Ann Steuve, Kathleen Gerson, and Lynne McCallister Jones, with Mark Baldassare. 1977. *Networks and Places.* New York: Free Press.

Freedman, Jonathan. 1975. *Crowding and Behavior.* San Francisco: Freeman.

Freeman, Linton C 1976. *A Bibliography of Social Networks* Exchange Bibliographies nos. 1170–71. Monticello, Ill.: Council of Planning Librarians.

Friedmann, Harriet. 1974. "Are Distributions Really Structures? A Critique of the Methodology of Max Weber." Research paper no. 63. Toronto: Centre for Urban and Community Studies, University of Toronto.

Gans, Herbert. 1962. *The Urban Villagers.* New York: Free Press.

———. 1967. *The Levittowners.* New York: Pantheon.

Gates, Albert S., Harvey Stevens, and Barry Wellman. 1973. "What Makes a Good Neighbor?" Paper presented at the annual meeting of the American Sociological Association, New York, August.

Gillies, Marion, and Barry Wellman. 1968. "East York: A Profile." Mimeographed. Toronto: Clarke Institute of Psychiatry, Community Studies Section.

Gordon, Michael. 1976 "Kinship Boundaries and Kinship Knowledge in Urban Ireland." Mimeographed. Storrs: University of Connecticut, Department of Sociology.

———. 1977. "Primary Group Differentiation in Urban Ireland." *Social Forces* 55 (March): 743–52.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (May) · 1360–80.

———. 1974. *Getting a Job* Cambridge, Mass : Harvard University Press.

———. 1976. "Network Sampling: Some First Steps." *American Journal of Sociology* 81 (May): 1287–1303.

Greer, Scott. 1962. *The Emerging City* New York · Free Press

Gurr, Ted Robert. 1969. *Why Men Rebel.* Princeton, N.J.: Princeton University Press.

Gusfield, Joseph R. 1975 *Community: A Critical Response.* New York: Harper & Row.

Howard, Leslie 1974. "Industrialization and Community in Chotanagpur." Ph D dissertation, Harvard University

Hunter, Albert. 1975 "The Loss of Community: An Empirical Test through Replication " *American Sociological Review* 40 (October) · 537–52

Jacobs, Jane 1961 *The Death and Life of Great American Cities.* New York · Random House.

Jacobson, David. 1973. *Itinerant Townsmen: Friendship and Social Order in Urban Uganda* Menlo Park, Calif : Cummings.

———. 1975. "Fair-Weather Friend." *Journal of Anthropological Research* 31 (Autumn): 225–34.

Janowitz, Morris 1952 *The Community Press in an Urban Setting.* Glencoe, Ill.: Free Press.

Kadushin, Charles. 1966. "The Friends and Supporters of Psychotherapy: On Social Circles in Urban Life " *American Sociological Review* 31 (December): 786–802.

Kasarda, John D , and Morris Janowitz 1974 "Community Attachment in Mass Society " *American Sociological Review* 39 (June): 328–39.

Katz, Fred 1966. "Social Participation and Social Structure." *Social Forces* 45 (December): 199–210.

Katz, Pearl 1974 "Acculturation and Social Networks of American Immigrants in Israel." Ph D. dissertation, State University of New York at Buffalo

Keller, Suzanne 1968 *The Urban Neighborhood.* New York Random House.

Kemper, Theodore D. 1972. "The Division of Labor: A Post-Durkheimian Analytical View." *American Sociological Review* 37 (December): 739–53.

Klatzky, Sheila R. 1971. *Patterns of Contact with Relatives* Washington, D C.: American Sociological Association.

Klovdahl, Alden S 1977. "Social Networks: Selected References for Course Design and Research Planning " Mimeographed. Canberra: Australian National University, Department of Sociology.

Kornhauser, William. 1968. "Mass Society." Pp 58–64 in *International Encyclopedia of the Social Sciences.* Vol. 10. New York: Macmillan and Free Press.

Laslett, Peter. 1971. *The World We Have Lost* 2d ed. London · Methuen.

Laumann, Edward O. 1973. *Bonds of Pluralism.* New York: Wiley.

Laumann, Edward O , Joseph Galaskiewicz, and Peter Marsden. 1978. "Community Structures as Interorganizational Linkages." *Annual Review of Sociology* 4: 455–84.

Lee, Nancy Howell. 1969. *The Search for an Abortionist.* Chicago: University of Chicago Press.

Leyton, Elliott, ed. 1974. *The Compact: Selected Dimensions of Friendship.* Newfoundland Social and Economic Papers, no. 3. St. John's: Institute of Social and Economic Research, Memorial University of Newfoundland.

Liebow, Elliot. 1967. *Tally's Corner.* Boston: Little, Brown.

Litwak, Eugene. 1960. "Occupational Mobility and Extended Family Cohesion." *American Sociological Review* 25 (February): 9–21.

Litwak, Eugene, and Ivan Szelenyi. 1969. "Primary Group Structures and Their Functions." *American Sociological Review* 35 (August): 465–81.

Marx, Leo. 1964. *The Machine in the Garden.* New York: Oxford University Press.

Mayer, Philip, and Iona Mayer. 1974. *Townsmen or Tribesmen.* 2d ed. Capetown: Oxford University Press.

Meier, Richard L. 1968. "The Metropolis as a Transaction-maximizing System." *Daedalus* 97 (Fall): 1293–1313.
Merton, Robert. 1957 "Patterns of Influence: Local and Cosmopolitan Influentials." Pp. 387–420 in *Social Theory and Social Structure*. Rev. ed. Glencoe, Ill.: Free Press.
Mitchell, J. Clyde. 1969. "The Concept and Use of Social Networks." Pp. 1–50 in *Social Networks in Urban Situations*, edited by J. Clyde Mitchell. Manchester: University of Manchester Press.
———. 1974. "Social Networks." *Annual Review of Anthropology* 3· 279–99.
Mostacci, Livianna. 1976 "The Degree of Multiplicity of Relationships in the Urban Setting." Typescript. Toronto: University of Toronto, Department of Sociology.
Nadel, S. F. 1957. *The Theory of Social Structure*. London: Cohen & West.
Nisbet, Robert, 1969. *The Quest for Community* New York: Oxford University Press
Park, Robert E. 1925*a* "The City: Suggestions for the Investigation of Human Behavior in the Urban Environment." Pp. 1–46 in *The City*, edited by Robert E. Park, Ernest W. Burgess, and Roderick D McKenzie. Chicago: University of Chicago Press.
———. (1925*b*) 1967 "The Urban Community as a Spatial Pattern and a Moral Order." Pp. 55–68 in *Robert E. Park on Social Control and Collective Behavior*, edited by Ralph H. Turner. Chicago: University of Chicago Press.
Pickvance, C. G. 1975. "Voluntary Associations and the Persistence of Multiple Ties." Mimeographed. Manchester: University of Manchester, Department of Sociology.
Rapoport, Anatol, and William J. Horvath. 1961. "A Study of a Large Sociogram." *Behavioral Science* 6: 279–91.
Ross, Mark Howard, and Thomas S. Weisner 1977. "The Rural-Urban Migrant Network in Kenya." *American Ethnologist* 4 (May): 359–75.
Rytina, Steve. 1977. "Sampling and Social Networks: Relational Measures in a Small Population" Working Paper no. 166. Ann Arbor: Center for Research on Social Organization, University of Michigan.
Scott, Joan, and Louise Tilly. 1975. "Women's Work and the Family in Nineteenth-Century Europe" *Comparative Studies in Society and History* 17 (January): 36–64.
Shorter, Edward. 1975 *The Making of the Modern Family*. New York: Basic.
Shulman, Norman. 1972. "Urban Social Networks." Ph.D. dissertation, University of Toronto.
———. 1976. "Network Analysis: A New Addition to an Old Bag of Tricks." *Acta Sociologica* 19 (December): 307–23.
Simmel, Georg (1902–3) 1950. "The Metropolis and Mental Life." Pp. 409–24 in *The Sociology of Georg Simmel*, edited and translated by Kurt Wolff. Glencoe, Ill.: Free Press.
———. (1908) 1971. "Group Expansion and the Development of Individuality." Pp. 251–93 in *Georg Simmel: On Individuality and Social Forms*, edited by Donald N. Levine and translated by Richard P. Albares. Chicago: University of Chicago Press.
Stack, Carol B. 1974. *All Our Kin*. New York: Harper & Row.
Stein, Maurice. 1960. *The Eclipse of Community*. Princeton, N.J.: Princeton University Press.
Suttles, Gerald D. 1972. *The Social Construction of Communities*. Chicago: University of Chicago Press
Taub, Richard P., George P. Surgeon, Sara Lindholm, Phyllis Betts Otti, and Amy Bridges. 1977. "Urban Voluntary Organizations, Locality Based and Externally Induced." *American Journal of Sociology* 83 (September): 425–42.
Tilly, Charles. 1970. "Community: City: Urbanization." Rev. version. Ann Arbor: University of Michigan, Department of Sociology.
———. 1973. "Do Communities Act?" *Sociological Inquiry* 43 (December): 209–40.
———. 1975. "Food Supply and Public Order in Modern Europe." Pp. 380–455 in *The Formation of National States in Western Europe*, edited by Charles Tilly. Princeton, N.J.: Princeton University Press.

———. 1978. *From Mobilization to Revolution* Reading, Mass. Addison-Wesley.

Tönnies, Ferdinand. (1887) 1955. *Community and Association*. Translated by Charles P. Loomis London: Routledge & Kegan Paul

Valentine, Charles. 1968 *Culture and Poverty*. Chicago· University of Chicago Press.

Verbrugge, Lois M. 1977. "Multiplexity in Adult Friendships" Rev version of paper presented at the 1976 annual meeting of the American Sociological Association, New York, September.

Walker, Gerald. 1977. "Social Networks and Territory in a Commuter Village, Bond Head, Ontario." *Canadian Geographer* 21 (Winter): 329-50.

Walker, Robert L. 1974. "Social and Spatial Constraints in the Development and Functioning of Social Networks: A Case Study of Guildford" Ph D dissertation, London School of Economics.

Warren, Donald I., and Rachelle B. Warren. 1976. "The Helping Role of Neighbors: Some Empirical Findings" Mimeographed. Detroit· Oakland University, Department of Sociology.

Warren, Roland L. 1978. *The Community in America* 3d ed Chicago Rand McNally

Wayne, Jack. 1971. "Networks of Informal Participation in a Suburban Context." Ph D. dissertation, University of Toronto

Webber, Melvin 1963 "Order in Diversity: Community without Propinquity," Pp 23-54 in *Cities and Space: The Future Use of Urban Land*, edited by Lowdon Wingo, Jr. Baltimore: Johns Hopkins Press.

Weisner, Thomas S. 1973 "The Primary Sampling Unit· A Nongeographical Based Rural-Urban Example" *Ethos* 1 (Winter)· 546-59.

Wellman, Barry. 1976. "Urban Connections." Research Paper no 84. Toronto· Centre for Urban and Community Studies, University of Toronto.

———. 1977. "The Community Question" Research Paper no. 90. Toronto: Centre for Urban and Community Studies, University of Toronto.

Wellman, Barry, Paul Craven, Marilyn Whitaker, Sheila Dutoit, and Harvey Stevens. 1971. "The Uses of Community" Research Paper no 47 Toronto: Centre for Urban and Community Studies, University of Toronto

Wellman, Barry, and Barry Leighton 1979. "Networks, Neighborhoods and Communities: Approaches to the Study of the Community Question" *Urban Affairs Quarterly* 15 (March): in press.

Wellman, Barry, Norman Shulman, Jack Wayne, and Barry Leighton In preparation *Personal Communities in the City*. New York Oxford University Press

Wellman, Barry, and Marilyn Whitaker, eds. 1974. *Community-Network-Communication: An Annotated Bibliography*. Bibliographic Paper no 4. 2d ed Toronto: Centre for Urban and Community Studies, University of Toronto.

White, Harrison C. 1965. "Notes on the Constituents of Social Structure." Mimeographed. Cambridge, Mass.· Harvard University, Department of Social Relations

———. 1966. "Coupling and Decoupling." Mimeographed. Cambridge, Mass: Harvard University, Department of Social Relations.

White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions." *American Journal of Sociology* 81 (January): 730-80.

White, Morton, and Lucia White. 1962. *The Intellectuals versus the City*. Cambridge, Mass.: Harvard University Press.

Wireman, Peggy. 1978. "Intimate Secondary Relations" Paper presented at the Ninth World Congress of Sociology, Uppsala, August.

Wirth, Louis. 1938. "Urbanism as a Way of Life." *American Journal of Sociology* 44 (July): 3-24.

Wolf, Eric B. 1966. "Kinship, Friendship and Patron-Client Relations in Complex Societies." Pp. 1-22 in *The Social Anthropology of Complex Societies*, edited by Michael Banton. London: Tavistock.

Woodsworth, J. S. (1911) 1972. *My Neighbor*. Toronto: University of Toronto Press.

Young, Michael, and Peter Willmott. 1957. *Family and Kinship in East London*. London: Routledge & Kegan Paul.

Differences in the Canadian and American Class Vote: Fact or Pseudofact?[1]

John F. Myles
Carleton University

This paper reassesses Alford's comparative analyses of class voting in Canada and the United States and updates them to 1976. It is argued that Alford's finding that the class vote in Canada is substantially lower than that in the United States is an artifact of his classification scheme, which incorrectly superimposes the American party structure on Canadian political parties. When Canadian political parties are classified according to the major party/third party division which more accurately reflects the realities of the "democratic class struggle" in Canada and when the correct techniques for measuring class voting are applied, it is shown that class voting in the two countries is virtually identical. It is not the level of class voting which distinguishes the two societies but rather the differential manner in which the democratic class struggle in each has been organized as a result of their different electoral systems.

The traditionally weak link between social class and electoral politics in the United States has made the "American exception" the subject of considerable speculation and empirical research among analysts of the "democratic class struggle" in Western societies. However, since the appearance of Alford's (1963, 1967) seminal analyses of class voting in the Anglo-American democracies, it has been accepted that while class voting in the United States is low, in Canada it is almost nonexistent. This paper will demonstrate that the apparently more marked "Canadian exception" is a pseudofact resulting from both theoretical and methodological problems associated with the analysis of class voting in Canada.

On theoretical grounds, there is little reason to expect a large difference in the class vote in the two countries since, as Alford himself noted (1963, p. 101), Canada and the United States are remarkably similar with respect to those factors usually called upon to explain a low level of class struggle. Where the two countries do differ, however, is the manner in which their respective electoral systems have mediated and organized this struggle. In

[1] Numerous colleagues and friends provided helpful criticism and comments on earlier drafts of this paper. In particular, I should like to thank Barry George, Robert Hauser, Hugh McRoberts, Rick Ogmundson, Dennis Olsen, and John Zipp. I am also grateful to the staff of the Social Science Data Archive of Carleton University for their assistance in making data for this paper available to me.

the United States, class-based protest movements have largely been absorbed within the existing two-party structure. In particular, the Democrats, under Roosevelt, successfully became identified as the representatives of labor and hence as the party of the left. In Canada, on the other hand, no such alignment of a major party with working-class interests occurred. The Canadian parliamentary system has discouraged the absorption of class and sectional protest movements into the two major parties while actively encouraging their expression in the formation and institutionalization of third parties (Lipset 1976). The result is that Alford's classification of the Liberals and the New Democratic Party (until 1961 the Cooperative Commonwealth Federation [CCF/NDP]) on the left and Conservatives and Social Credit on the right is highly misleading. The left-right division of the two major parties, Liberals and Conservatives, in a manner analogous to the Democrat-Republican split in the United States, is particularly problematic.

In his review of the problem of classifying Canadian political parties, Ogmundson (1975) notes that the original classification proposed by Alford does not correspond to voter perceptions of the class interests represented by the different parties, their bases of electoral support, the social origins of their elected members, or their sources of party financing. On all of these indicators, the two major parties (Liberals and Conservatives) are clearly representative of the middle class while the two minor parties (CCF/NDP and Social Credit) are representative of the working class. Accordingly, he argues that estimation of the class vote in Canada should be based on a classification which distinguishes between the two major parties on the one hand and the two minor parties on the other. This problem of conceptual equivalence, however, is further compounded by the procedure used to measure the class vote.

MEASURING THE CLASS VOTE

Alford defines the class vote as the ". . . extent to which manual and nonmanual strata divide in their support for political parties" (1963, p. 79). His index of class voting is computed by ". . . subtracting the percentage of persons in non-manual occupations voting for Left parties from the percentage of persons in manual occupations voting for Left parties" (1963, pp. 79–80). Alford's index is, in fact, identical with Somer's *d*, which in turn is identical with the slope of the regression of vote on class in the 2×2 table (see Korpi 1972, p. 631). This use of the regression approach when the dependent variable is a dichotomy has been criticized for various reasons (Goodman 1972; Theil 1970). For our purposes, the major difficulty is that the index is sensitive to the marginal distributions which generate the cell values and, hence, potentially confounds the effect of class

on vote within societies with differences in the overall level of right or left voting between societies. Such a property is undesirable since, as Alford observed, ". . . it is the gap between the voting patterns of manual and non-manual occupations which is vital for assessing class voting, not the over-all level of Right or Left voting" (1963, p. 80). Given the distributions in Alford's original analysis, this difficulty did not actually arise. However, when the classification of parties suggested above for Canada is employed, it becomes severe since the distribution of the dependent variable is highly skewed.

A strategy which overcomes the difficulties inherent in the usual regression approach is the use of odds ratio based measures which are insensitive to variations in the marginals. Several such approaches are available, each of which will provide equivalent results. One such method is the application of a modified multiple regression procedure based on log-linear techniques, as recently elaborated by Goodman (1972). An equivalent solution, which provides results more readily comparable with Alford's analysis, is to remove the effects due to marginal distributions by means of Deming adjustment (Deming 1943). This procedure involves the progressive standardization of rows and columns to a common set of marginals for all tables. It preserves the association between row and column variables (the relative odds ratio) while removing the differences due to the effects of the column and row marginals (Mosteller 1968). Following Mosteller's (1968) example, each 2 × 2 table (vote X class) included in the analysis was standardized on row and column marginals set equal to 100, and a standardized index was then computed in the usual way. This procedure was applied to all except two of the election polls originally utilized by Alford (1963, appendix B) as well as to two polls conducted in conjunction with each national election in Canada and the United States since 1961.[2]

RESULTS

The results from the reanalysis of Alford's (1963) data and the results for subsequent federal elections in Canada and the United States are presented in the Appendix. Inspection of the adjusted indices indicates a similar pattern for the two countries. The more recent results are summarized in table 1, which replicates and updates Alford's (1963, p. 102) summary table 5-2. Two sets of measures are presented for Canada: those computed when Alford's original classification of parties is used (Canada 1) and those computed using the major party/third party split (Canada 2). In each case two values are presented: the adjusted index of class voting

[2] The two polls excluded are American polls for 1940 and 1944 (sample 1) for which sample sizes were not provided and hence cell values could not be computed. Details on the selection of polls are available in the lengthier version of this paper available on request from the author.

TABLE 1

CLASS VOTING IN CANADA AND THE UNITED STATES, 1952–61 AND 1962–76

	1952–61						1962–76					
	Mean		Lowest		Highest		Mean		Lowest		Highest	
	A^*	A	A^*	A	A^*	A	A^*	A	A^*	A	A^*	A
United States	17	16	14	13	24	23	15	14	6	5	20	18
Canada 1	8	8	−2	−1	18	17	1	1	−6	−6	10	9
Canada 2	17	9	10	5	26	12	16	11	10	2	22	18

NOTE.—A^* = adjusted index; A = unadjusted index, Canada 1 = CCF/NDP and Liberals classified as left, Canada 2 = CCF/NDP and Social Credit classified as left.

after differences due to row and column marginals are removed (denoted by A^*) and the original indices as computed by Alford (denoted by A).

For the period 1952–61, it is evident from an examination of rows 1 and 2 that standardization has little effect on Alford's original indices. However, examination of row 3 indicates that when the major party/third party split is employed the unstandardized indices seriously underestimate the effect of class on vote in Canada. Moreover, it is clear that the size of the class vote in Canada and the United States is virtually identical over this time period. The mean for both countries is identical, while the range in Canada is slightly greater than that in the United States at both extremes.

For the more recent period, the consequences of reclassifying the Canadian parties are even more striking. Using Alford's original classification (Canada 1), the average class vote for Canada is almost zero, reflecting the fluctuation between low positive scores and low negative scores which appear in the Canadian data. One could only conclude from such results that class voting has declined in Canada over the last decade and a half to the point of being nonexistent. The results from the major party/third party classification (Canada 2), however, indicate a very different picture. Rather than declining, the level of class voting in Canada has remained constant. Once again, the Canadian average for the period is virtually identical with that for the United States.

CONCLUSION

In both Canada and the United States, the representatives of the working class are constrained in their pursuit of pure class political goals. Within the American context, the labor wing of the Democratic party is but one faction among several which must be appeased through coalition politics. In Canada, third-party representatives of the working class must rest content with the role of critic and proponent of policies which may or may not be adopted by the party in power. Whether coalition or third-party politics provides a more effective means of achieving (or suppressing) working-class interests is a topic which lies beyond the scope of this paper. For purposes of comparative political analysis, however, it is clear that future research should address the North American exception and that the Canadian exception is a pseudofact which can be laid to rest.

APPENDIX—INDICES OF CLASS VOTING IN CANADA AND THE UNITED STATES, 1936–76

YEAR	*A** Canada 1	*A** Canada 2	*A** United States	*A* Canada 1	*A* Canada 2	*A* United States
1936			16, 16			16, 15
1940	8	18		6	6	
1944			14, 21			14, 21
1945	6, 8	22, 16		6, 8	16, 12	
1948			20, 30, 43			20, 29, 41
1949	4	30		4	17	
1951	10	28		9	15	
1952			24, 18			23, 17
1953	8	18		8	10	
1954	8	20		8	12	
1956	2	14	14, 16	2	8	13, 16
1957	10, 10	12, 16		9, 10	6, 8	
1958	16	26		16	9	
1959	−2	16		−1	8	
1960	6	16	16, 14	7	8	15, 14
1961	18, 4, 8	22, 20, 10		17, 4, 8	12, 12, 5	
1962	2, 6	22, 14		1, 6	18, 12	
1963	−4, −6	16, 16		−3, −6	13, 12	
1964			16, 20			14, 17
1965	4, 0	10, 13		3, 1	2, 12	
1968	−6, 2	18, 20	18, 14	−5, 2	12, 14	17, 13
1972	4, −6	12, 8	6, 16	3, −5	9, 7	5, 15
1974	10, 2	16, 22		9, 2	11, 14	
1976			18, 14			18, 14

SOURCES.—Indices for 1936–61 recomputed from Alford 1963, appendix B Indices for 1962–76 computed from the following sources. Canada 1962 Canadian Institute of Public Opinion (CIPO) 296, 297, 1963 CIPO 301, 302; 1965 CIPO 315, 1965 Canadian National Election Study; 1968 CIPO 329, 330, 1972 CIPO 355, 356; and 1974 CIPO 366, 367. United States 1964 American Institute of Public Opinion (AIPO) 698, 699; 1968· AIPO 770, SRC 1968 National Election Study; 1972 AIPO 858, 859, and 1976 AIPO 959, 961

NOTE.—*A** = adjusted index, *A* = unadjusted index, Canada 1 = CCF/NDP and Liberals classified as left; Canada 2 = CCF/NDP and Social Credit classified as left.

REFERENCES

Alford, Robert R. 1963. *Party and Society: The Anglo-American Democracies.* Westport, Conn.: Greenwood.

———. 1967. "Class Voting in the Anglo-American Political Systems." Pp. 67–93 in *Party Systems and Voter Alignments,* edited by S. M. Lipset and Stein Rokkan. New York: Free Press.

Deming, W. Edwards. 1943. *The Statistical Adjustment of Data.* New York: Wiley.

Goodman, Leo A. 1972. "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables." *American Sociological Review* 37 (February): 28–46.

Korpi, Walter 1972. "Some Problems in the Measurement of Class Voting." *American Journal of Sociology* 78 (November): 627–42.

Lipset, Seymour Martin. 1976. "Radicalism in North America: A Comparative View of Party Systems in the United States and Canada." Paper presented at the Joint Meeting of the Royal Society of Canada and the American Academy of Arts and Sciences, Laval University, Quebec City, June 7, 1976.

Mosteller, Frederick. 1968. "Association and Estimation in Contingency Tables." *Journal of the American Statistical Association* 63 (March): 1–28.

Ogmundson, Rick. 1975. "On the Measurement of Party Class Positions: The Case of Canadian Federal Political Parties." *Canadian Review of Sociology and Anthropology* 12 (November): 565–76.

Theil, H. 1970. "On the Estimation of Relationships Involving Qualitative Variables." *American Journal of Sociology* 76 (July): 103–54.

Commentary and Debate

The comments printed in this section are limited to brief critiques of articles and reviews published in the *AJS*. These comments are expected to address specific errors or flaws. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Authors of articles and reviews are invited to reply to comments, subject to the same length limits. The *AJS* does not publish commenters' rebuttals to authors' replies. Longer or less narrowly focused comments on articles published in the *AJS* should be submitted themselves as articles.

COMMENT ON LENSKI'S "HISTORY AND SOCIAL CHANGE"

In his article (*AJS* 82 [November 1976]: 548–64), Gerhard Lenski criticizes Robert Nisbet's *Social Change and History* (1969) for providing a distorted view of neoevolutionism. Lenski's critique must be faulted on three grounds: (1) his failure to recognize the underlying assumption behind Nisbet's criticism of neoevolutionism, (2) his failure to provide evidence for his own tenuous characterization of the new evolutionists, and (3) his unwarranted claim that the distinction between general and specific evolution should quell most criticisms of the neoevolutionary perspective.

While in *Social Change and History* Nisbet focuses primarily on the "fallacies" of all developmental theories, he also criticizes functional theorizing and what he calls the "conventional wisdom regarding change" in the social sciences. The social sciences are shown to embrace all or most of the central premises of the classical social evolutionary perspective (1969, pp. 235, 270), which he identifies as the following: (1) change is normal, (2) change is directional, (3) change is immanent, (4) change is continuous and cumulative, (5) change is necessary, and (6) change proceeds from uniform causes (1969, pp. 166–82). After identifying these premises, Nisbet proposes six diametrically opposed alternative principles that he regards as more faithful descriptions of empirical reality. Accord-

ingly, the first of the evolutionist premises is discarded in favor of one that states that fixity, not change, is normal and natural (1969, p. 270). The second premise, in turn, is repostulated from "change is directional" to "change lacks directionality" (1969, pp. 284–87), and so on.

Lenski comments disparagingly on Nisbet's "fondness for dichotomies" (p. 553) but fails to discern that Nisbet perceives a common assumption in all developmental theories—the assumption of a compelling determinism. Each of the premises of social evolution detailed by Nisbet reveals this common deterministic framework. For change to be defined as "directional," for example, it is logically imperative that change must be seen as inevitably constrained or made to conform to some natural or destined course of development. Similarly, the proposition that "change is immanent," or proceeds from some internal dynamic, virtually precludes the accidental, the random, the willfully intentional from having significant impact.

For Nisbet, then, the theme of necessity or determinism constitutes the overriding and most objectionable element in neoevolutionism. This is apparent from his criticism of the fourth premise—"change is continuous." He notes that the "real objection . . . is to the fixed notion within the conventional wisdom of social science that one change necessarily engenders another, that one 'stage' of developmental change produces the next stage . . . " (1969, p. 290). The primary shortcoming of neoevolutionism, he tells us repeatedly, is that its concerns with theory, natural history, natural law, and progress leave little room for the multitude of essentially indeterminate developments—"random events," "unique occurrences," "the purely fortuitous" (1969, p. 238), the "willful" and the "adventitious" (1969, p. 171).

Nisbet's commitment, therefore, to an approach in the social sciences that does not preclude on theoretical grounds the importance of indeterminate developments—a more historical approach—accounts for his categorical rejection of neoevolutionism. The question needs to be asked, however, whether the new evolutionism does in fact, as Nisbet claims, subscribe to a fundamental determinism. Lenski would have us believe that Nisbet's depreciation of the new evolutionism is explainable, in part, by his failure to recognize some existing differences between the old and the new evolutionism: "Most [contemporary evolutionists] would not claim that there is necessity in most patterns of change. . . . Most of us would probably be content to say that there are simply varying degrees of probability. One of the important differences between the 19th- and 20th-century forms of evolutionism is the shift from deterministic to probabilistic formulations" (p. 557).

First, who is included in "most of us"? Certainly Sahlins and Service's *Evolution and Culture* (1970), which Lenski cites approvingly, firmly

establishes their perspective as ". . . plain old evolutionary, not neo-evolutionary" (p. 4). And Leslie White, whose concept of evolution as an information-building process is regarded by Lenski as the "single master trend" (Lenski 1976, p. 555) in all of evolution, has this to say: ". . . Neo-evolutionism is a misleading term: it has been used to imply that the theory of evolution today is somehow different from the theory of eighty years ago. We reject any such notion" (1959, p. ix). Indeed, the single reference we are provided with is "Lenski and Lenski 1974, pp. 108–11" (p. 557), in which we encounter solely the authors' argument on behalf of probabilism, followed—without foundation—by the summary statement that "modern evolutionary theory does *not* take a deterministic view . . . " (1974, p. 110).

Second, what is behind the assertion that these unidentified modern evolutionists *"would probably be content* to say that there are *varying degrees* of probability" (italics added)? If there exists no underlying determinism, then why is there an apparent need to compromise—to settle for or "be content" with "varying degrees of probability"? What original position is it that they "would probably" compromise? In summary, the tenuous qualifiers "most of us" and "would probably be content" can hardly be seen as justifying Lenski's sweeping conclusion that a defining characteristic of 20th-century evolutionism is a probabilistic orientation.

But let us grant, for the sake of the argument, that some neoevolutionists have shifted to probabilistic formulations of their theories. What are the theoretical and practical implications of such a shift? For it is only from the nature of these implications that we may judge whether this alleged shift constitutes, as the author claims, an important difference.

It is Lenski and Lenski's *Human Societies* (1974) that provides an insight into the importance of this theoretical shift. "Evolutionary theory," they argue here, "does not assume that technological advance has been inevitable everywhere and all the time. . . . Technostasis and even technological regression have occurred frequently, and evolutionary theory must take account of this" (1974, p. 84). To do so, the authors invoke Sahlins's distinction between general and specific evolution. "General evolution is the evolution of human societies as a whole the world over; specific evolution is the evolution of an individual society. . . . [T]echnostasis and regression are much more relevant for specific evolution, since they have been infrequent phenomena on the global level but fairly common at the level of individual society" (1974, p. 84). As Lenski observes in his article, "Statements that are nonsensical when applied to some instance of the latter [specific evolution] may be entirely valid on the level of general evolution. . . . Many of Nisbet's objections to the newer evolutionism seem to stem from his failure to appreciate the extent to which it has focused on general evolution" (p. 562).

To Lenski the distinction between the two types of evolution enables social evolutionists to address a variety of criticisms. It provides an explanation for such phenomena as the fall of civilizations—phenomena described, significantly enough, as "deviations from the hypothesized pattern of growth and development" (p. 551). It is important, also, that this distinction is used to support the claim that the new evolutionism is probabilistic in its orientation, since it has ceased, according to Lenski, to regard progress in terms of determinism or absolute necessity. It is apparent, however, that Lenski's probabilistic stance remains confined, essentially, to individual case studies, or to the level of specific evolution. In that realm where evolutionists, by the author's admission, have focused their attention—namely, the realm of general evolution—the "shift" away from determinism is far from apparent. In fact, Lenski himself insists that detractors of neoevolutionism must make their case against determinism: "Thus, to disprove the theory that there is long-run directionality . . . critics must show that at the *global level* there is no trend with respect to central tendencies. . . . This, of course, cannot be done . . . " (p. 552).

Lenski's assertion, therefore, that neoevolutionists develop their theoretical models in probabilistic rather than deterministic terms cannot be taken at face value. In the area of macro changes social evolutionists consistently point to the existence of some natural or destined course of development, generally expressed in terms of increasing populations, social complexity, and so on.

According to the evolutionist perspective, there are forces that compel sociocultural change to stay on course and sustain the emergence of "superior forms." It is this recognition, one might speculate, that has led White, Service, and Sahlins to insist that their brands of evolutionism are indeed not neoevolutionism, but in the tradition of their 19th-century forerunners—and, therefore, clearly deterministic. Only by defining sustained directionality as probabilistic does one have a probabilistic rather than a deterministic model of change. If Lenski accepts this definition, however, his theory ceases to be what is, in any conventional use of the term, an evolutionary model.

In summary, then, Nisbet's underlying assumption—that the premises of neoevolutionism are grounded in determinism—appears to be fundamentally correct. Contrary to Lenski's assessment, the distinction between general and specific evolution does not make for a more probabilistic evolutionism. The difficulty with Lenski's brand of neoevolutionist position is that, at the primary level of explanation, it remains the evolutionism of the 19th century. Nisbet understands this; Lenski, apparently, fails to do so, and the nature of his attack reflects this failure.

George Becker

Vanderbilt University

REFERENCES

Lenski, Gerhard, and Jean Lenski. 1974. *Human Societies*. 2d ed. New York: McGraw-Hill.

Nisbet, Robert. 1969. *Social Change and History*. New York: Oxford Galaxy.

Sahlins, Marshall, and Elman Service, eds. 1970. *Evolution and Culture*. Ann Arbor: University of Michigan Press.

White, Leslie. 1959. *The Evolution of Culture*. New York: McGraw-Hill.

PROBABILISM REASSERTED: A REPLY TO BECKER

It is not easy to respond to Becker's comments, since he attacks me for a position which I explicitly rejected, not only in my paper "History and Social Change" but also in *Human Societies,* which he cites. If I read Becker correctly, his primary concern is with "the assumption of a compelling determinism," which he, like Nisbet, claims to find in the new evolutionary theory (and in other developmental theories and in functionalism as well). This is the common denominator underlying all three of Becker's criticisms of my paper.

I am glad for this opportunity to respond to Becker, both to set the record straight and to comment on the issue of determinism, which is an important one. For the record, in "History and Social Change" I began my discussion of the question of determinism by saying, "Probably the least objectionable of Nisbet's proposals concerning the nature of change is his assertion that '. . . [T]here is no necessity in any pattern of change if our attention is fixed, not upon the wholes and universals of social evolutionary theoretical interest, but instead upon the social behavior of human beings in time and place.' " In commenting on this, I said, "Actually, most contemporary evolutionists would be even more cautious than Nisbet in one respect. Most would not claim that there is necessity in most patterns of change—*even when attention is fixed on 'the wholes and universals of social evolutionary theoretical interest'* " (italics added). I went on to say, in the phrase Becker chooses to quote, "Most of us would probably be content to say that there are simply varying degrees of probability" (p. 557).

There are two points to be noted here. First, the italicized phrase makes it clear that my commitment to probabilism applies to general evolution, no less than to specific. Second, "the original position" which is being modified (not "compromised," as Becker would have it) is that of none other than Robert Nisbet, who implies a deterministic position regarding general evolution in the statement of his which I quoted. Perhaps Becker

is right that I should have been less conciliatory in my critique of (Nisbet's) determinism at this point.

Lest it seem that these few phrases from my recent paper represent an effort to cover up a more basic commitment to determinism, I would point to two features of the original edition of *Human Societies* (1970), which have been essentially preserved in the second and third editions (1974, 1978). In discussing the rise of the new evolutionary theory in biology, I noted that it involved a shift from deterministic to probabilistic formulations (pp. 59–60) and later suggested that evolutionary theory in sociology should follow this lead (p. 70). Still later, in the closing chapter of the book, I had the following to say in a discussion of the problems of predicting the future in a section headed "Determinism versus Probabilism Once Again": "Earlier in this volume, when we examined the issue raised by deterministic theories, we observed that most social scientists now avoid both extreme positions. They reject the determinists' thesis that the future course of events is inevitable or predetermined, and hence completely predictable if one only has sufficient information on relevant sociological variables. Similarly, they reject the thesis that it is impossible to make any predictions at all. Rather, they take a position that has come to be known as probabilism, maintaining that all *possible* future developments are not equally *probable*" (p. 498). I do not know how I can make my position any clearer.

I must confess, however, that I am troubled when I find intelligent sociologists, such as Nisbet and Becker, repeating again and again the charge that evolutionists are determinists. Why do they do it? And, above all, why do they do it in the face of clear evidence to the contrary and the disavowals of those they attack?

One possibility is that they assume mistakenly that a concern for trends that are sustained for centuries and millennia implies a belief in historical inevitability. Another possibility is that they are humanists who mistakenly, again, assume that any search for regularities in human life, whether by evolutionists or others (and there are clear indications in both Nisbet and Becker that their criticisms are not limited to evolutionists), is a threat to belief in freedom of the will. Or, finally, given their fondness for dichotomies, it is possible that they do not really see any meaningful distinction between probabilism and determinism, since both imply constraints on human freedom.

I cannot help wishing that Becker and others who share his perspective would respond to the assertion of some of us that there are important identifiable patterns in the historical record. What do they make of these patterns? Why do they avoid a discussion of them? Do they deny their existence? If not, how do they explain them? Do they really want us to believe that these long-term trends are merely the result of the fortuitous

and the adventitious, and that somehow or other purely random events fall out in such a way, century after century and millennium after millennium, as to sustain these trends? And, if not that, do they propose that sociology should ignore such trends? Answers to these questions might go a long way toward clarifying some of the confusion which they have introduced.

Before closing, I would like to respond to Becker's question as to who the other recent neoevolutionists are who share my nondeterminist perspective. To begin with, I would never include Leslie White among them, partly because he retired a good number of years ago and is now dead, and partly because his rigid determinism acted as a powerful negative model for me, and probably for others as well. White had some important contributions to make, not least of which was his role as a transitional figure between the older and newer evolutionism, but his position on the issue of determinism has been abandoned, to the best of my knowledge, even by his former students, Service and Sahlins.

More to the point, when I read the work of people such as Boulding (1964, 1970, 1978), Parsons (1966, 1971, 1977), Bellah (1964), or Winch (1978), I find very little evidence of a deterministic perspective. Probabilism, yes; determinism, no. The very use of contingency tables (note the adjective!) by Winch and others of us constitutes an implied commitment to probabilism.

Let me close this statement as I closed my article on Nisbet, with a plea for an end to the quarrel between humanistic historians, with their focus on unique events and unique personalities, and evolutionists, with their focus on the regularities and patterned trends which emerge from the study of the historical record. There is a need for both in today's world. Humanistic history is a valuable antidote to fatalism, just as evolutionism is a check on utopianism. The former reminds us that the human will can influence events, just as the latter reminds us that human will always operates within the framework of external constraints. So let us not subvert the study of the role of the contingent in history by denying its existence, and, similarly, let us not deny the need for the study of historical patterns and trends by refusing to acknowledge their presence.

GERHARD LENSKI

University of North Carolina at Chapel Hill

REFERENCES

Bellah, Robert. 1964. "Religious Evolution." *American Sociological Review* 29 (June): 358–74.

Boulding, Kenneth. 1964. *The Meaning of the Twentieth Century: The Great Transition.* New York: Harper & Row.

———. 1970. *A Primer on Social Dynamics: History as Dialectics and Development.* New York: Free Press.

———. 1978. *Ecodynamics: A New Theory of Social Evolution.* Beverly Hills, Calif.: Sageview.

Lenski, Gerhard. 1970. *Human Societies* New York: McGraw-Hill.
Parsons, Talcott. 1966 *Societies: Evolutionary and Comparative Perspectives*. Englewood Cliffs, N J · Prentice-Hall
———. 1971. *The System of Modern Societies* Englewood Cliffs, N.J.: Prentice-Hall
———. 1977 *The Evolution of Societies* Englewood Cliffs, N J : Prentice-Hall.
Winch, Robert. 1978 *Familial Organization*. New York: Free Press

COMMENT ON MOLLENKOPF'S REVIEW

"While New York's importance may be declining in some respects . . . , it may actually be increasing in other respects" (p. 475), writes John Mollenkopf in his recent book review (*AJS* 83 [September 1977]: 474–79). If so, we would have a long overdue corrective to the widely held opinion among urbanologists and regional scientists that trends in population and employment are a robust indicator of a region's overall state of economic health. "In fact, declining populations may even be functional" (p. 475). Urbanologists need to be reminded that the "nerve center" of a nation does not have to endlessly accumulate population and capital from the remainder to retain its central role in that nation's economy, and that in fact it need not have very large size at all—the activities in the hinterland that are coordinated by the center may in certain periods call for a disproportionate amount of the nation's energy and physical investment, but the center could remain nonetheless primary for this growth and vigor in the peripheral parts. After all, the brain of a human roughly doubles in size by maturity, while the body reaches 20 times its original size; but the brain retains its central control function despite its slower growth rate. Mollenkopf cites some surprising statistics on trips and messages to and from New York City that seem to demonstrate New York's continued vigor as the nation's nerve center despite its loss of population: "Airline passenger figures into New York have risen steadily, and the Manhattan exchanges are still unexcelled for number of interstate phone calls" (p. 475).

The published statistics on the flow of messages and trips between the various parts of this country that I am familiar with, however, give a rather different impression. Passengers enplaned in the three major airports of the New York metropolitan area have not only not risen steadily (they declined in 1974 and, along with the remainder of the country, during the recession of 1975), but the number enplaned in 1976 is an insignificant 0.2% above the number enplaned in 1973. In contrast, the number enplaned in the country as a whole grew by 10% between 1973 and 1976 (table 1). I know of no published data that give the number of interstate calls processed by the Manhattan exchanges. But the Federal

TABLE 1

TOTAL PASSENGERS ENPLANED

	New York*	(% change)	U S. Total	(% change)
1973	17,639,108		202,208,000	
1976	17,674,566	(+.2)	223,313,000	(+10 4)

SOURCE.—Air Transport Association of America, *Air Transport 1974* (Washington, D.C.. Air Transport Association of America, 1974), pp 3, 34; *Air Transport 1977*, pp. 5, 8.

* La Guardia, John F. Kennedy, and Newark airports.

TABLE 2

NUMBER OF CALLS ORIGINATING FROM TELEPHONES (in Thousands)

	Local	(% Change)	Toll	(% Change)
	NEW YORK STATE			
1970	18,021,597		1,296,834	
1975	19,144,800	(6 2)	1,288,688	(−.6)
	U S TOTAL			
1970	169,110,360		10,611,420	
1975 ..	216,889,110	(28 3)	17,913,387	(68 8)

SOURCE.—Federal Communications Commission, *Statistics of Communications Common Carriers 1970* (Washington, D C. Government Printing Office, 1972), p. 16, table 9, *Statistics of Communications Common Carriers 1975*, p 13, table 8.

Communications Commission does publish data on the number of calls originating in each state on an annual basis. These data show a decline in toll calls originating in New York State between 1970 and 1975 (table 2). In the nation as a whole, toll calls increased by more than two-thirds over this same five-year period. The trends in New York City should track those of New York State.

Until Mollenkopf reveals data that tell us otherwise, students of this nation's cities and regions have no real choice but to stick with their dull but serviceable "one-factor" model of regional economic activity. Not only are population and capital draining off from the center into the periphery, but there is no evidence that higher-level control functions are not as well. (There are a few data points: In 1977, 83 *Fortune* 500 companies have their head offices in the so-called Sunbelt, compared with 59 in 1970 [*Fortune* (June 1977): 144].) Mollenkopf's thesis that the regional hierarchy has remained stable over the 1970s despite radical changes in the relative growth of its different levels is an interesting and imaginative one, but there is little in the data with which I am familiar to make me confident that it has any basis in reality.

DANIEL R. VINING, JR.

University of Pennsylvania

REPLY TO VINING

Daniel Vining has elegantly summarized my argument that the center can continue to coordinate activities in the hinterland even while the latter grows more rapidly. He contends, however, that two indexes which I argue reflect roughly the strength of this coordination role—namely, passengers boarding at New York airports and phone calls made from the city's exchanges—do not support my thesis. There are two issues here: we need to know the actual facts of the matter with regard to plane passengers and phone calls, and we need to specify more accurately what it means for a big city to retain and benefit from a higher-order, coordinating, or controlling role in the urban hierarchy. Although I believe the facts support my original contention, whether they do or not may be irrelevant if the original hypotheses have been improperly framed.

First, the facts. Vining has chosen an aberrant period in which to examine trends, namely, the serious national economic downturn which began in 1972. Since we want to know about New York's business relations, a severe recession is probably not a good time to examine them. And, indeed, if one looks instead at the longer and more representative period between 1965 and 1976, one finds that New York's passengers have increased 70% while the nation's grew 147%. More important, if one examines the networks among cities, they remain reasonably stable with New York at the center. This pattern changes mainly as cities which appear themselves "tributary" to New York extend their own penetration into smaller cities.

Phone calls are more complicated. New York State figures contain many stray influences other than New York City's business influence. (Only about half the state's phones are in the city.) And out-of-state phone calls, even more than plane boardings, reflect a broader set of motives than business-related activities alone. Finally, leased lines and computer-driven communications have totally revolutionized the meaning of a phone call over the past decade. Even so, a spokesman for the AT & T long-lines division assured me off the record that their New York City equipment is being constantly upgraded to respond to increasing demand. Interestingly, AT & T declines to publish data on point-to-point communications patterns for fear of alerting competition to potential markets. This suggests that a look at what AT & T's competitors are trying to do would reveal New York's importance. And, indeed, the first competing transmission system is a new tie between New York and Houston. While this suggests Houston's growing importance, it certainly does not suggest that New York is on the decline. In short, to the extent that we can make inferences from longer-term trends in these two admittedly crude indicators, they

support rather than "flatly contradict" the argument as initially stated.

This leads us to the second issue: Aggregate statistics about individual cities and their relative rates of growth and decline are to some degree beside the point, because they do not measure the interactions between cities very well. The essence of the "retention of control" hypothesis is that actors located in New York play important roles in things that happen in other cities, and, conversely, that actors in the other cities cannot decide their affairs among themselves alone but must somehow involve New York actors. Direct evidence on such exchanges across cities (whether they take place in the public or the private sector) would be vastly superior to indirect indicators such as New York's aggregate volume of telephone calls. Unfortunately, few data of this sort are published. If any reader has detailed information of this type, whether for phone networks, bank clearings, or whatever, all parties would undoubtedly be glad to hear about it.

Such crude information as is available also supports the retained-influence hypothesis for New York. Although *Fortune* 500 firms have moved away from Manhattan, and although the number located in the Sunbelt has increased, the New York metropolitan area has actually retained the same number of headquarters over the past decade or so. And this fact should not be particularly surprising, given that New York has retained its role at the top of the hierarchy through many periods of substantial social change. (For an interesting study of the stability of the urban hierarchy, see Madden [1955/56].)

To conclude, the empirical indicators challenged by Vining support my case, as does other admittedly partial and indirect evidence. Yet he has done a service by pointing out that we need to have a much more specific investigation of the question, both in terms of how the relevant hypotheses are framed and what type of evidence is gathered. A definitive resolution clearly awaits further research.

John Mollenkopf

Stanford University

REFERENCES

Madden, C. H. 1955/56. "On Some Indications of Stability in the Growth of Cities in the United States." *Economic Development and Cultural Change* 4:236–52.

Review Essay: On Tourism and Modern Culture[1]

The Tourist: A New Theory of the Leisure Class. By Dean MacCannell. New York: Schocken Books, 1976. Pp. xi+214. $12.95 (cloth); $5.95 (paper).

Image and Pilgrimage in Christian Culture: Anthropological Perspectives. By Victor Turner and Edith Turner. New York: Columbia University Press, 1978. Pp. xvii+281. $15.00.

Hosts and Guests: The Anthropology of Tourism. Edited by Valene L. Smith. Philadelphia: University of Pennsylvania Press, 1977. Pp. 254. $14.00.

Michael S. Schudson
University of Chicago

Freud constructed psychoanalytic theory out of attention to the trivial occurrences of everyday life, such as dreams and slips of the tongue. Dean MacCannell tries to reconstruct social theory by attending to a commonplace escape from everyday life—tourism.

MacCannell would not be surprised at the comparison with Freud. As he undertakes his ethnography of tourism, he argues that ethnography "will eventually occupy a position in the modern world similar to the one occupied by psychoanalysis in the industrial world" (p. 174). The modern or postindustrial world, in MacCannell's view, is the "coming to consciousness of industrial society" (p. 182). It is marked, in the world of science, by sociology, and in the everyday world by tourism. MacCannell argues that the tourists frequently surpass the sociologists in understanding the meanings of modern life. And well they might, for, if MacCannell is right, tourist attractions provide the language through which people in modern society seek an authentic understanding of their world.

This notion may sound a little loony. *The Tourist* has an unflattering habit of announcing revelations rather than sharing information. It is pockmarked with ex cathedra pronouncements ("the courtroom is the most important institution in a democratic society") and epigrammatic misrepresentations of social theory ("Marx and Mao are committed to the priority of the material substratum over theory and ideas"). Still, it is a serious stab at, all at once, creating useful categories for the study of tourism and other forms of leisure; exploring the nature of authenticity in modern life; and seeking to construct grounds for an ethnography of complex society, a way

[1] I want to thank John MacAloon, Carl Pletsch, Mike Ryan, and Robert Snow for stimulating my thinking on tourism.

of comprehending the signs by which a highly differentiated society is given meaning. The book is not what the subtitle says it is, a "new theory of the leisure class." One of its more egregious faults is that it says nothing about class at all. MacCannell is "suspicious of research that insists on the primacy and independence of social class" (p. 11). This is scant justification for ignoring class entirely. But if *The Tourist* is not a theory of the leisure class, it *is* a theory of leisure. And it is, warts and all, a dazzling book.

There are good grounds for sociological attention to tourism. Take simply the place of tourism in the world economy. In 1950, international tourism alone was a $2 billion business. In 1960, it accounted for $6.8 billion; 1970, $17.9 billion; 1975, $31.9 billion. The number of international tourists has grown enormously. In the United Kingdom, for example, there were just over 600,000 tourists in 1950, 1.7 million in 1960, and 5.9 million in 1970. Overall, there were 72 million international tourists in 1961 and 168 million in 1970. In 1965, international tourist receipts represented 3.4% of total exports for the United States, 14.3% for Greece, 19.0% for Ireland, 39.4% for Mexico, and 47.7% for Spain. While Europe and North America still attract the most tourists, Third World countries have actively promoted tourism, too, with the United Nations, the World Bank, and other aid agencies assisting them.[2]

That tourism is of great and growing economic significance has made it politically important. Governments, both national and local, are involved in promoting and in regulating tourism. Whereas the creation of tourist attractions (like Disney World) has been the work of entrepreneurs and developers, the preservation of sights of possible touristic interest has concerned a wide variety of individuals and groups, including community organizations, universities, banks, and city-planning commissions. Historical preservation has become a major consideration in the past several decades and frequently takes the shape of a social movement. Neighborhoods not only restore themselves physically by an appeal to tourism, but also reawaken culturally. Whole regions and whole countries may do so, too.

All this, however, while clearly of concern for the sociology of economic development, urban sociology, and the study of social movements, is not what MacCannell proposes as the central sociological question about tourism. He believes that tourism provides the symbols by which modern people give their lives meaning. Ours is not a world of religion or even of civil religion, in his perspective, but of tourism. The travel brochure, not the religious tract, is a source of inspiration. While, as I shall indicate, MacCannell takes this too far, I think he is right to suggest that a sociology of culture, broadly conceived, should pay attention to tourism, that the roadside historical markers we pass by or, more rarely, stop at, are signs of how we conceive our society and ourselves.

There is little in the classical tradition of either sociology or anthropology to help us study tourism. One thinks, perhaps, of the literature on "marginal men." But the marginal man stands outside or next to the mainstream

[2] For statistics on tourism, see Gray (1970), Popovic (1972), British Tourist Authority (1971), and World Bank (1972).

of society; the tourist stands above it. The stranger lives as a stranger; the tourist *looks* as a stranger but lives somewhere else.

One may turn to the sociology of religion for guidance. There is a difference, of course, between a "holy day" and a "holiday," between a vacation "taken" or "spent" and a holy day "observed" or "celebrated." But MacCannell argues that we must understand tourism as religious ritual. According to him, the tourist is the model for "modern-man-in-general," and tourism is one of the two poles of modern consciousness. The other pole, revolution, is the desire to transform things; tourism is the inclination to venerate things as they are. Tourism is a ritual activity. Tourist attractions are "precisely analogous to the religious symbolism of primitive peoples," and sightseeing can be defined as "a ritual performed to the differentiation of society." Modernity, for MacCannell, *is* structural differentiation, the source of both the sense of freedom in modern life and the sense of alienation, fragmentation, and discontinuity. Tourism explores, venerates, studies, and shares in structural differentiation.

That is MacCannell's—I think largely mistaken—claim. Just what does it mean? Fortunately, MacCannell is forthcoming with illustrations drawn from tour books and other travel literature and from his own field notes as a participant-observer tourist. The object of tourism is society itself. Tourism tries to see society and grasp its connections in an enjoyable way. Workers, alienated from their labor in the factory, may nonetheless tour some other place of work, the FBI or the mint or another factory, and enjoy watching the work that they would not enjoy doing. Tourist "sights" may include all the different features of society. One may tour neighborhoods, work districts (the garment district in New York), domestic establishments ("Society Hill" in Philadelphia or antebellum mansions in Natchez), groups (Pennsylvania Dutch or Indians on reservations), occupations (blue-collar work in a factory, journalism in a newsroom, farming at a dairy), transportation (the Golden Gate Bridge), and public works (MacCannell cites a tour of London's public lavatories). MacCannell discusses turn-of-the-century guidebooks to Paris, including Baedeker's, which he calls "still the best ethnography of modern cities." The Paris guidebooks direct the visitor to and give detailed descriptions of the stock exchange, the morgue, the supreme court, the mint, the government printing office, the Gobelin tapestry weavers, a tobacco factory, a slaughterhouse, and the sewers. In the tourist's guidebook, experience is democratized. Everything deserves attention. Just as, in Durkheim's view, primitive peoples unknowingly worshiped the power of society itself in their totems, so, in MacCannell's perspective, tourists sacralize the variety of social formations, the sum of which represents the forces of society.

For MacCannell, however, not all tourist attractions are equivalent. They differ according to their degree of authenticity. Religion has to do with being in the presence of ultimate reality and feeling subject to the power of the ultimate. Tourism, in its own fashion, is a quest for the ultimately real or, as MacCannell puts it, for the authentic in human experience. MacCannell orders tourist settings along a continuum of authenticity. Some attrac-

tions are purely "front," set up solely to attract tourists. Others are purely "back," not designed for tourists at all but come upon only by chance or persistence—these are "unmarked" attractions. In between, there are fronts organized to look like backs (a pioneer village like Old Sturbridge), backs open to tourists (factory tours), or backs spruced up a bit for visitors (the open orchestra rehearsal). The whole range of attractions plays on the modern passion for experiencing reality and participating in the authentic. In a marvelous aperçu which dramatizes this argument, MacCannell writes that the person who wants to get off the beaten path, who disdains commercialization and the most superficial tourist shrines, is not the antitourist but the supreme tourist: ". . . touristic shame is not based on being a tourist but on not being tourist enough, on a failure to see everything the way it 'ought' to be seen. The touristic critique of tourism is based on a desire to go beyond the other 'mere' tourists to a more profound appreciation of society and culture, and it is by no means limited to intellectual statements. All tourists desire this deeper involvement with society and culture to some degree; it is a basic component of their motivation to travel" (p. 10). Seeking the authentic, which means avoiding commercialization, avoiding other tourists, and even avoiding consciousness of oneself as a tourist, is the essence of tourism.

That is MacCannell's argument, always provocative, sometimes brilliant, and, I think, wrong. Is it true that tourists seek authenticity? Some do, some do not, and the tourist industry caters to both kinds. Large numbers of tourists, even when traveling abroad, seek food they recognize, music they are familiar with, and comforts to which they are accustomed. H. Peter Gray distinguishes two forms of tourism—"wanderlust," whose aim is education, and "sunlust," whose aim is rest (Gray 1970, p. 14). This suggests correctly that, contrary to MacCannell, tourism is not a system of attractions. Instead, it is a system of attractions and *dis*tractions. Tourism is as much, or more, a moving away as a moving to. It is escape as often as involvement.

In this sense, tourism is not "precisely analogous to the religious symbolism of primitive peoples." MacCannell himself observes that religious shrines in modern society lose their traditional value and take on touristic significance: "Throughout the world, churches, cathedrals, mosques, and temples are being converted from religious to touristic functions" (p. 43). But this admits that modern ritual (which is not religious) differs from primitive ritual (which is). MacCannell's assertion that tourism is ritual is too glib. So is his view that tourism has moral significance—after all, he asks, do we not speak of a tourist attraction as a "must see"? But if you pass up a "must see," you do not feel you have transgressed divine or social codes. You have simply failed to do right by yourself, failed to experience something you might have taken in. MacCannell makes a contribution in urging us to attend to the ways in which tourism is a crusade for authenticity, but he is misleading in suggesting that this is all of tourism or even the essence of tourism. What may be more essentially touristic is not a search for the authentic but a doubt that it can be found. One feels more

sure that a touristic journey will provide distraction than meaningful attraction, and tourism lives in the space between those poles.

Even this, however, gives MacCannell's case too much credit, for it assumes that we are all modern and, as moderns, engage in tourism. But this is not so. In any given year, in the United States, most people are tourists, if at all, only in their own communities. In 1972, 45% of Americans (or 37% of American households) did not take a single trip to a destination 100 miles or more away from home. Among those who did travel, 39% of their trips were to visit friends and relatives, 20% were for business and conventions. Freud seems to have been right—love and work are what matter or, at least, what still move people the most. Only 12% of trips were for outdoor recreation and 13% for sightseeing and entertainment (U.S. Bureau of the Census 1973).[3] Of course, these data are not conclusive, but they at least suggest that people organize the meanings of their lives in a circle centered on home and family, work and community, rather than in a far-flung system of tourist shrines.

When tourism does seem most surely a quest for authenticity, is it in fact a ritual "performed to the differentiation of society"? I think not. On the contrary, when tourism is most nearly religious ritual, it is an escape from social differentiation, an effort to transcend the limits of daily social experience by coming into contact with nature, or with culture in the form of unifying symbols and myths which transcend social differentiation. MacCannell acknowledges that attractions at the heart of tourism may escape commercialization, and here he refers to natural sights. The Grand Canyon, despite tours, hotels, postcards, and the Chamber of Commerce, is not reproducible. It can be experienced in only one place in the world. MacCannell finds it significant that California, a state rich in natural wonders, should be forty-seventh in the amount of state funds spent to promote tourism. Why is there such a persistent popularity of natural sights if tourism is a worship of social differentiation?

On the other side, if tourism worships differentiation, what accounts for the popularity of sights of cultural unity? Why is tourism the second largest industry in Washington, D.C.? Why do millions visit Lenin's tomb every year?

There is, I think, an answer to these questions. It is implicit in a fascinating new book by Victor Turner and Edith Turner, *Image and Pilgrimage in Christian Culture*. The Turners discuss the symbols and social structure of Christian pilgrimage, covering both medieval pilgrimages and, in greater detail, contemporary pilgrimages in Mexico, Ireland, and France. Their main interest is "the theological doctrines and popular notions which promote and sustain Christian pilgrimage, and the symbols and images which embody them." Much of the book, then, a rather fine-grained anthropological account of particular symbolic structures, is not related to the concerns of this review, but some of it can be read as part of a conversation with students of tourism as a social and cultural phenomenon. The Turners

[3] In addition, 16% traveled for "other" reasons.

can be seen to offer both significant support and serious critique of the kind of position MacCannell takes.

The support is simply this: the Turners suggest that tourism is worth serious attention from social scientists. In their chapter on medieval pilgrimage, they indicate that the serious meaning of a pilgrimage was often covered over by the barnacles of a tourist trade. Pilgrim ways in the Middle Ages were lined with way stations that displayed sensational relics, many of which, like the trinkets of tourism in a later age, assumed the values originally attached to the place to which the pilgrimage was made. The relics became fetishes "operating by principles of sympathetic and contagious magic, rather than serving as vehicles of religious and ethical ideas" (p. 197). Bordellos were easily available along pilgrimage routes. Adultery was rumored to be common at pilgrim shrines. Pilgrimage tended to lose its penitential quality and to become "worldly and fashionable." Reformers and satirists attacked pilgrimage and "began to raise a cry which became louder and louder until it was something like a storm at the time of the Reformation" (p. 184).

Despite the corruptions of pilgrimage, it played (and still plays today) a vital cultural role in two respects. First, it creates what the Turners call a socioeconomic "field." In the Middle Ages, pilgrimages stimulated and shaped the growth of cities, marketing systems, and roads—perhaps as much, the Turners suggest, as purely economic and political considerations.

Second, pilgrimage did, for many, and still does today, provide people with a sense of relationship to ultimate values. Here, I think, is the kernel of a critique of MacCannell. The Turners suggest that pilgrimage, far from being a worship of society, is a transcendence of the social, an experience that, drawing from the language of Victor Turner's earlier works, lifts people out of *societas* or "structure" and into *communitas* or "antistructure." In new terminology borrowed from Mihaly Csikszentmihalyi's (1975) studies on play and games, the Turners argue that the meaning of pilgrimage lies in the experience of "flow" it engenders. Flow is "the merging of action and awareness, the crucial component of enjoyment." Flow, experienced in play, sport, artistic performance, and religious ritual, is the product of centering attention on a limited field, assisted often by a set of rules. We then act with total involvement, aware but not aware of our awareness: "There is a loss of ego, the self becomes irrelevant." The Turners suggest that pilgrimage has survived the centuries and, in one form or another, may be regarded as a cultural universal, because it has regularly succeeded in producing this overwhelming feeling of enjoyment, "flow."

Is there a sense of "flow" in tourism? When tourism is most like pilgrimage, when it is "some form of deliberate travel to a far place intimately associated with the deepest, most cherished, axiomatic values of the traveler," there may well be a sense of flow. Pilgrimage, plain and simple, is still a major element of modern travel, as the Turners show. Tourist bureaus and travel agencies must be well aware that 700,000 people annually travel to Knock in Ireland, 3 million a year go to Lourdes, 3 million to the cathedral of Santiago de Compostela in Spain, and more than 5 million to the shrine

of Guadalupe in Mexico. Secular travel which is most like pilgrimage, I would suggest, is that which draws people out of social diversity into nature or into symbols of national or international unity. Contrary to MacCannell, the authentic attractions people seek are not things that are genuinely what they are but things that are *more* than what they are. The Lincoln Memorial is accordingly richer than the mint or the FBI. Indeed, the machine or tractor displayed at the world's fair, where it represents national power and solidarity, has a richness of association greater than the same machine or tractor displayed to tourists in its more "authentic" setting on factory or farm.[4]

MacCannell's book can be contrasted to a recent collection of papers by anthropologists, *Hosts and Guests: The Anthropology of Tourism*, edited by Valene Smith. *Hosts and Guests* is a very different, and lesser, effort than MacCannell's at understanding tourism. Smith et al. look at the effects of tourism on small-scale societies, while MacCannell attends to the meaning of tourism in modern societies. Smith et al. concentrate on the people who become tourist attractions or workers in the tourist trade, while MacCannell focuses on the people through whom they make their living, middle-class tourists. MacCannell's is a book of great pretension, Smith's one of modesty. It includes short pieces on Eskimo tourism and on the manufacture of *molas* by the San Blas Cuna Indians of Panama. In his contribution on tourism in Tonga, Charles F. Urbanowicz tells us that the Tongan Visitors Bureau found tourists disappointed that Tongans knew little of their own culture and history, so the bureau published a pamphlet for local consumption to improve on this score. A piece on Bali suggests that tourism may "strengthen the process of conserving, reforming, and recreating certain traditions." On the other hand, Davydd Greenwood documents how tourism changed the *Alarde*, a major public ritual in the Spanish Basque country, from a vital commemoration of a victory over the French in 1638 to an obligation to be avoided. The municipal government urged that the ritual be given twice in the same day, redefining it as a "public show to be performed for outsiders." One of the best pieces in the volume is Oriol Pi-Sunyer's discussion of native-tourist interaction in Catalonia, an essay which neatly undercuts the volume's "hosts and guests" title. For, as Pi-Sunyer argues, the native-tourist relationship is the opposite of the traditional host-guest relationship of Catalonians: "If tourism commoditicizes culture, natives categorize strangers as a resource or a nuisance rather than as people" (p. 155).

Unfortunately, the essays are afterthoughts to anthropological fieldwork, not genuine case studies. In few of them do we hear native voices. This is not the "anthropology of tourism" it hopes to be—actually, MacCannell's study of guidebooks comes closer to that, closer to investigating tourism itself as a cultural system. Of course, the anthropologists have no more resources for dealing with tourism in their tradition than do the sociologists.

[4] Attractions of cultural unity are sometimes less compelling than people most eager to stimulate cultural unity would like. Philadelphia anticipated 14–45 million visitors during the bicentennial and got 7–10 million instead As Michael Zuckerman laments, "Surveys showed that travelers were celebrating the bicentennial in California, the Southwest, and Hawaii instead" (Zuckerman 1978, p. 225).

As Smith observes, anthropologists have been slow to address themselves to tourism. The first published article was in 1963; the first national symposium, which generated *Hosts and Guests,* was in 1974.

One result is a failure to find a consistent tone or stance regarding tourism. Some of the essays try to be analytic, some are oriented to questions of policy ("Tonga dances have a beauty and symmetry of their own that should be performed for the visitors to Tonga"), and some sound like cliché-ridden copy for the tourist industry itself ("North Carolina is rich in natural resources, and her coastal area is unrivaled for sheer beauty anywhere on this continent. Her greatest wealth, however, lies in the people of North Carolina"). A question the anthropologists are all concerned with is, Is tourism good or bad for local people in developing parts of the world?

They awkwardly avoid an answer. The tone of the book conveys a rather studied refusal to be judgmental. The authors sound as if they are trying to convince themselves that tourism is potentially a good thing. Theron Nuñez, in the concluding essay, notes the tendency of anthropologists to be hostile to the very notion of tourism. While urging anthropologists to overcome this attitude, Nuñez reveals his own allegiance to it: "At the present state of our knowledge it is difficult to demonstrate that tourism *per se* is uniquely destructive or evil." Smith similarly fights her instincts in order to say something nice about tourism. While noting unhappily that tourism leads to the transfer of local control to national governments and the domination of local trade by international business, she argues still that "the tourist trade does not have to be culturally damaging." She believes that a little tourism is a good thing but that "mass tourism" is likely to dehumanize both hosts and guests. Indeed, she distinguishes several types of tourism according to the number of tourists and their degree of adaptation to local norms. The greater the number of tourists (as in "mass" tourism or "charter" tourism), the poorer the adaptation; the smaller the number of tourists ("explorer" tourism or "elite" tourism), the better the adaptation.[5]

This stance is exactly what MacCannell, in his final chapter, calls the "antitourist" position. Antitourism, he suggests, is a position of modern urban liberals and Third World radicals which "can be reduced to a negative response to the recent expansion of the tourist class to incorporate larger numbers of different kinds of people" (p. 164). Antitourists enjoy their own tourism but oppose "these other tourists." MacCannell's own position is that both protourists, who favor tourism for economic development (despite evidence that tourism can help "underdevelop" a nation by keeping capital in foreign hands and providing primarily menial employment), and antitourists are responding to the growth of tourism which has outstripped "support" institutions. He sees a solution in increasing "the number of marked attractions and support facilities for tourists on a worldwide base." But, he acknowledges, this may not work. To skip industrialization and

[5] There is surely a need for more attention to the social costs, and benefits, of tourism. For other comment on the question, see Bryden (1973, pp. 91–96) and Wood (1979); see also Wood's unpublished paper, "International Tourism and Cultural Change in Southeast Asia."

jump headfirst into modernity through becoming a tourist "attraction" is a dangerous path for a nation. "If a group avoids industrialization as it modernizes, it may end up spending all of its time wondering if it is real or not—that seems to be the price paid for rapid advancement in the modern world."

That may also be the price paid for a book, like MacCannell's, which playfully leapfrogs over Marx and Veblen, pokes at Lévi-Strauss and Goffman and Benjamin (while taking some hints from them), and absurdly claims to have transformed the study of modern culture from "theoretical speculation" to "empirical studies" (p. 110). Foolhardy it is to claim to be doing an ethnography of tourism without a word about the age, sex, class, nationality, or personality of tourists, without a word on the daily life of tourism, without mention of the social organization of the support institutions of hotels, airlines, and tour companies. (And, of course, they are not only support institutions, but institutions which invent new attractions to support themselves.) MacCannell does not come to grips with the problem posed by the importance of natural sights, nor does he even acknowledge the problem posed by sights of cultural integration for his theory. All this leads the reader to wonder if *The Tourist* is real or not. It has, perhaps, advanced too rapidly, from unquestionable insights to theoretical parody.

One early tourist MacCannell quotes as having written in 1876 (Arthur Young) in fact wrote in 1790, a time when the usual tourist was still the aristocrat on the grand tour, an eighteen-year-old equipped with guidebooks, a pistol, and a telescope and often accompanied by a tutor. Even at that time tourists could be nuisances. Voltaire, at his estate at Ferney, found himself too often a tourist attraction, and once snapped at visitors who, unwelcomed, were staring at him: "Well, gentlemen, you now see me. Did you take me for a wild beast or monster that was fit only to be stared at as a show?" (Hibbert 1969). It is not a fair question. If tourists keep their distance from some attractions, it may be as much from humility as from arrogance. It is a ritual distance, if MacCannell is right. It is an "observant" stare. What this does to celebrities, who become sights, is worth considering in the same breath with what it does to the people who live in tourist neighborhoods or to the people of the Bahamas, where nearly two-thirds of employment is in the tourist trade.

If Daniel Bell is right that culture has come unhinged from social structure, if culture now leads rather than follows social change, so consciousness, as MacCannell suggests, may not be an outgrowth of structure so much as the definition of what social structure is today. Alternatively, if the Turners are right, then culture has long been a template for social change. People's need to participate in a meaningful cosmos has had quite as much to do with the patterns of settlement and commerce as have more strictly economic needs. The economy of the soul as well as the economy of the body must be satisfied. Indeed, the Turners speculate, the seeds of new social structures may be sown in old antistructures, and medieval pilgrimage may have been "a prototype, under different socioeconomic conditions, of more egalitarian, voluntaristic, contractual types of secular social relationships."

In any case, the problem MacCannell started with is a serious one: Is an ethnography of modern, differentiated, and worldwide integrated society possible? Is there any way to capture the meaning of modern experience? At least, as MacCannell begins to show, and as the Turners show in focusing on contemporary mass phenomena more closely linked to traditional religion, we can examine the ways in which we and others seek for meaning. We can examine the experiences of pilgrimage and tourism, the ways in which people gather meanings at their shrines, and capture them for future reference in memory, in photographs and films, or, for that matter, in monographs.

REFERENCES

British Tourist Authority. 1971. *Digest of Tourist Statistics No. 2* London: British Tourist Authority.

Bryden, John M. 1973. *Tourism and Development.* Cambridge: Cambridge University Press.

Csikszentmihalyi, Mihaly. 1975 *Beyond Boredom and Anxiety: The Experience of Play in Work and Games.* San Francisco: Jossey-Bass

Gray, H. Peter. 1970. *International Travel—International Trade.* Lexington, Mass.: Heath.

Hibbert, Christopher. 1969. *The Grand Tour.* New York: Putnam.

Popovic, Vojislav. 1972 *Tourism in Eastern Africa.* Munich· Weltforum.

U.S. Bureau of the Census. 1973. *Census of Transportation, 1972.* Vol 1. *National Travel Survey* Washington, D C : Government Printing Office.

Wood, Robert E 1979. "Tourism and Underdevelopment in Southeast Asia." *Journal of Contemporary Asia,* vol. 8, no. 4.

World Bank 1972 *Tourism: Sector Working Paper* Washington, D.C : World Bank.

Zuckerman, Michael. 1978. "The Irrelevant Revolution· 1776 and Since" *American Quarterly* 30 (Summer): 224–42.

Review Essay: Feminism and Sociology—Some Recent Perspectives

Dependence and Exploitation in Work and Marriage. Edited by Diana Leonard Barker and Sheila Allen. London: Longman, 1976. Pp. viii+265 $15.50 (cloth); $7.95 (paper).

Sexual Divisions and Society: Process and Change. Edited by Diana Leonard Barker and Sheila Allen. London: Tavistock Publications, 1976. Pp. ix+287. £6.50 (cloth); £3.50 (paper).

Being Female: Reproduction, Power and Change. Edited by Dana Raphael. Chicago: Aldine Publishing Co., 1975. Pp. xiii+293. $14.50.

Woman, Culture and Society. Edited by Michelle Zimbalist Rosaldo and Louise Lamphere. Stanford, Calif.: Stanford University Press, 1974. Pp. xi+352. $12.50 (cloth); $4.75 (paper).

Ann Oakley
University of London

Sociological capital can be made out of social protest movements. Specific sociological products (books, articles, doctoral dissertations) are, of course, only one representation of the social resonance of protest. Moreover, sociological analysis is bound to annex protest to its own mainstream concerns; the task of social science, as many of its practitioners see it, is the intellectual colonization of social action. Yet certain brands of protest may be assimilated into prevailing modes of sociological (or anthropological) analysis only with great difficulty.

The relationship between sociology and feminism is a complicated one. Within the feminist movement sociologists are very often mistrusted because they are seen to turn women (like other social actors) into objects of study, whereas feminism decrees that women cease to be objects and become, properly, subjects of their own experiences. Marxist feminists further complain that sociologists implicitly support the exploitation of women because they are themselves agents of capitalist ideology (even if their work is not directly financed by a government or geared specifically to the indoctrination of future workers). Yet there exist, paradoxically, feminist sociologists—as the four volumes reviewed here evidence. This conjunction implies that sociology is not feminist; if this is so, feminists might reasonably inquire, what business do feminists have with it? Such an inquiry is double edged. In the first place, one wants to know in what ways sociology runs counter to feminist concerns. Second, the concern of feminism must be with the usefulness of sociology as one among many de-

vices for developing a concrete understanding of women's situation (historically, contemporarily, and cross culturally).

Each of these volumes is a collection of essays on different dimensions of being female and on the disjunctions and contradictions created by the attempt to marry conventional social science with feminism. There are five editors and 70 contributors; the subject matter of the collections, all published between 1974 and 1976, can be taken as one index of the manner in which social science, since the mid-1960s, has responded to the challenge of the women's movement. To list the subjects covered provides a rough guide to the focus of response. The largest number of pieces concerns motherhood and aspects of gender role differentiation (11 essays each); marriage is next (nine essays); followed by women's social networks, paid work, and domestic power (five each); then socialization (four); nondomestic politics, law, and the position of women in general (three each); and housework (two). The agenda thus implicitly disputes the importance, at this particular moment in our understanding of women's position, of looking at women's participation in formal institutional structures—politics, organizations, religion, and so on—and instead highlights the central significance of the domestic domain, informal structures, and general patterns of social division between the sexes in influencing the status of women. There is a concomitant stress on ideology—its meaning, creation, maintenance, and impact—at the expense of a materialist approach to the role of women in different cultural milieus.

Perhaps the most striking development in sociological literature on women reflected in these works is hidden by the colorful rhetoric of particular concerns. The onus is no longer (as it was a decade ago) on the feminist social scientist to prove the subordination of women in Western industrial society; this has become a taken-for-granted fact. Modern social institutions are guilty until proved otherwise, and it is up to the defenders of equality to identify nonsexist practices, beliefs, causes, and consequences as evidence against the (now implicit) allegation of patriarchy. (The word "oppression" is still, however, not much favored by social scientists, particularly in Britain; the American authors in these collections are more outspoken about their ideological commitment.)

Everyone in these volumes is agreed that in modern industrial society, at least (or especially?), women are regarded and treated as inferior, and that within the dominant ideology feminine concerns are trivialized and masculine ones inflated, so that a girl learns the devaluation of femaleness from the cradle to the grave. The most interesting pieces are those which combine original research with a critical attitude toward mainstream sociology's (or anthropology's) definition of woman's place. In their equitable preface to *Sexual Divisions and Society,* Diana Barker and Sheila Allen quote the following lines from Vanessa Maher's study of relationships among Moroccan women (prefaces in such collections are often the most fertile hunting ground for good ideas): " 'Only after a substantial, polemical, and varied experience of the society in question is it possible to dis-

cover that statements gravely and precisely articulated are statements of ideology, reflecting either the behaviour of the dominant social group, which may be the minority, or the ideological tendency to confuse the substance of an act with its legal enactment, thus attributing the agency of all significant social processes to men,' " (p. 3). This applies both to the common procedure among anthropologists of acquiring information from elderly, high-status males and to much of what has traditionally passed for social science. In her extremely thoughtful "Theoretical Overview" in *Woman, Culture and Society,* Michelle Rosaldo points out that much traditional ethnography, in describing the symbolic conceptions of woman's place that prevail within dominant masculine ideology, has ignored the fact that it is just these conceptions that may be used by women as a basis for female solidarity. Convents, brothels, informal neighborhood friendships, women's political organizations (in Africa, for example) are some instances of women's cooperation that coexist with an ideology that sets women apart and appears to constrain their activities. The culture of women is not necessarily identical with women's culture as portrayed by men, and, as certain of these authors make plain, special techniques are required by ethnographers from Western or other dominant cultures who are concerned with exposing the culture of a minority group. A logical consequence, which is barely touched on in any of these volumes, is that the analysis of women's culture as in some sense a counterculture to that of men implies the same masculine viewpoint that feminist social science is now challenging. So, for example, it is theoretically possible that rather than masculine conceptions of women providing a symbolic basis for women's consciousness and community, the existence of female solidarity precedes the erection by men of a symbolic account which allows them to feel they are controlling the place of women in society.

In what I suspect is, at times, an unthinking way, some of these pieces borrow the language of masculine institutions to describe female domains. Thus Ronald Frankenberg, in an amusing account of sexism in his own and other people's work on community, refers to "the industrial sociology of the housewife" (*Sexual Divisions and Society,* p. 28), and Ann Whitehead, in a pertinent analysis of female-male relationships in a Herefordshire village, talks of "the industrial sociology of child-rearing" (*Dependence and Exploitation,* p. 198). The use of such language certainly reflects the paucity of conceptual development as applied to the domestic world (and for this reason perhaps should be excused), but the danger is that, as with certain models of sex equality or liberation, we are simply fitting women to the contours of the masculine stereotype. This obfuscates the representation of a female perspective and in the long run is not the appropriate medicine with which to cure the inherited sexism of social science. As Rosaldo notes in her "Theoretical Overview," there are two arrangements which can potentially elevate women's status: either women may enter the public world (of men) or men may enter the domestic world (of women). But the formula needs a caveat, for societies tackling gender inequalities

have learned over the past 10 years that it is not enough to increase the participation of women outside the home. Individual women who benefit from expanded opportunities in the public domain may enjoy higher status, but the position of those who do not is entirely unaffected. The low status of domestic work and child rearing, caused by a continuing connection with devalued femininity, persists and is even intensified by the creation of two "classes" of women—those who can pretend to be like men and those who cannot.

The papers included in the two volumes edited by Barker and Allen were all given at the 1974 British Sociological Association conference entitled "Sexual Divisions and Society." This was the first formal sociological occasion in Britain which recognized gender roles as a serious object of study, one which was no longer to be swept under the carpet or dealt with in the odd paper or the occasional paragraph or two as "the problem of women." In this respect British feminist sociologists are slower on the uptake and less self-assertive than their American counterparts; Barker and Allen relegate to a footnote (*Sexual Divisions,* p. 18) their comments on the professional sociological (male) reaction to the idea of a conference on gender roles. I would have liked to see in print a detailed account of this historic occasion—an account which included all the backroom scenes (for example, the pressure from feminists outside sociology to choose a title more specifically focused on the position of women than the gentle and unexceptionable "Sexual Divisions").

One important task of feminist social scientists in our culture seems to me that of offering a continuing ethnography of the development of gender relations. To focus one's analysis on the situation of women rather than that of men (because women are the minority group) does not and should not preclude attention to the relational aspect of gender. A major theme emerging from these collections is the crucial impact of the context in which gender behavior is studied. Colin Bell and Howard Newby (*Dependence and Exploitation,* pp. 152–68) put the argument that women's deference is contextual. It is in the "dialectic" of marriage, where women feel themselves to be, are, and are expected to be relatively powerless in relation to men that feminine deference is most prominently displayed and most legitimated as a form of social (gender role) relationship. Much the same point is made by Marilyn Hoskins in an essay on the roles and options of Vietnamese women (*Being Female,* pp. 231–48). She draws a contrast between the formal respectful language and deferential attitude displayed by Vietnamese wives to their husbands and the role of many women in running the family business and handling the family income; both behavioral complexes are part of cultural womanhood in Vietnam.

Sensitivity to situational dimensions of gender role behavior follows from the growing input of research in this area. So, also, does a questioning of conventional research categories: marriage, work, socialization, politics, authority, and so forth. Increasingly it seems that the opposition between domestic and public domains in most, if not all, cultures is the main struc-

tural component of gender differentiation. Materially, socially, symbolically, and ideologically men and women are separated through their differential commitments in and outside the home. To assert this—as several contributors to these volumes do—clearly requires greater specificity about the meaning of "domestic." Just what is it that women are identified with, and—a different but essential question—why is it so?

In a fascinating chapter entitled "Family Structure and Feminine Personality" (*Woman, Culture and Society*, pp. 43–66), Nancy Chodorow puts forward the old argument that women's mothering role accounts for their universal subordination. However, Chodorow's proposal has none of the debased biologism of traditional polemics. Briefly, her contention is that because women give birth to children they are likely to play a larger part in their rearing than men; through early socialization by the mother, both female and male children experience the beginning of gender personality differentiation even before the onset of separate gender identities. (Thus Chodorow opposes the traditional psychoanalytic model in maintaining that the pre-oedipal experience is already different for boys and girls.) Feminine gender identity, on the one hand, is personal and consists of a long, unbroken, and, in some ways, unbreakable identification with the mother. This type of identification does not necessarily lead to personal adjustment since, Chodorow argues, the failure to separate from the mother is the source of many problems encountered by Western women (for instance, the fact that the biological activities of childbirth and breastfeeding, by challenging body boundaries, can be felt as intensely conflictual). Masculine gender identity, on the other hand, is positional, an achieved, learned status. To become a man a male child must repress femininity within himself and denigrate it outside: the achievement of masculinity is contingent on the subordination of women. One lesson that follows from this relates to the point I made earlier, that only a change in the definition of masculinity will affect the cultural devaluation of femininity. Men must accept "feminine" (domestic) interests as their own.

Thematically, the domestic-public duality is accompanied by an ideological opposition between nature and culture. This notion is very widespread in social representations of femininity. Sherry Ortner, in the chapter following Chodorow's ("Is Female to Male as Nature Is to Culture?") outlines one meaning of this: once again because of motherhood, women are easily linked with nonhuman nature. Women's biology encourages a definition of them as closer to nature than men. "Nature" becomes "women's nature." What is not feminine is seen (by men) as cultural, social order as opposed to natural disorder. The appropriation of social order by men as *the* social order, culture itself, further legitimates the "natural" inferiority of women.

Whichever way you look at it, it seems, having children is woman's essential handicap, her original sin. To claim that the universal subordination of women stems from childbearing is perhaps stretching a point, yet as a metaphor it usefully indicates the *difficulty* (but not impossibil-

ity) any culture has in overriding the tendency for the biological possession of a uterus to be culturally transformed into the possession of secondary social citizenship.

Moreover, it is clear from these volumes that the conventional sociological/anthropological focus on the institution of marriage per se has deflected interest from the even more powerful institution of motherhood. Christine Delphy, presenting a new view in her "Continuities and Discontinuities in Marriage and Divorce" (*Sexual Divisions and Society,* pp. 76–89), contends that it is because the family is the place where women rear children that marriage is "the institution by which gratuitous work is extracted from a particular category of the population, woman-wives" (p. 77). From children's dependence flows women's dependence and their "commitment" to marriage as the most valid social context in which dependence may be exercised. Deliberate efforts to promote the disintegration of conventional gender role behavior, as in some communes, stumble on the same cornerstone of motherhood, as Philip Abrams and Andrew McCulloch explain in a chapter entitled "Men, Women and Communes" (*Sexual Divisions and Society,* pp. 246–75): "What happens when the constraints of legally enforced monogamy and the wage-work–housework division of roles are removed, is not the creation of an alternative system of personal relationships, but rather the making plain of the deeper obstacles to equality through which men and women in this society struggle to relate" (p. 272).

If motherhood prevents equality, childbearing must be a burden. But—the usual question—whose voice is speaking here? Social and economic circumstances—isolation, poverty, lack of help—may make mothers feel worn down by their children, but the view of motherhood as uninteresting, marginal work is largely that of the male social scientist (a point made by Raphael on p. 66 of *Being Female* in a discussion of van Gennep's The *Rites of Passage*). To identify childbirth and childrearing as sources of inequality—like poverty or inadequate health care—does not mean that their eradication is a worthwhile or necessary goal. Many feminists (e.g., S. Firestone in *The Dialectic of Sex*) have collaborated with male ethnographers in equating motherhood with handicap and have outlined a vision of technological reproduction as the only true liberator of women (see H. Rose and J. Hanmer's "Women's Liberation, Reproduction and the Technological Fix," pp. 224–48 in *Sexual Divisions and Society*). And yet many mothers see their ability to give birth to, feed, and care for children as the greatest achievement of which they—or anyone else, female or male—are capable. (This insight can be extracted from many studies of marriage; see, for instance, L. Rainwater, L. P. Coleman, and G. Handel's *Workingman's Wife,* 1959; and H. Z. Lopata's *Occupation Housewife,* 1971. See also A. Oakley, *Becoming a Mother,* forthcoming.) To see childbearing as a resource rather than as biological and social weakness is a necessary correction of sexism both within and outside social science. The denial of all sex differences, part of the intellectual armory of femi-

nism, is a conceit of patriarchy, an overreaction of the underprivileged whose vision of themselves must necessarily undergo a long and painful, but ultimately fruitful, gestation.

REFERENCES

Firestone, S. 1972. *The Dialectic of Sex: The Case for Feminist Revolution.* London: Paladin.

Lopata, H. Z. 1971. *Occupation Housewife.* New York· Oxford University Press

Oakley, A. 1979. *Becoming a Mother.* London: Martin Robertson

Rainwater, L., L. P. Coleman, and G. Handel 1959 *Workingman's Wife: Her Personality, World and Life Style.* New York: Oceana.

Book Reviews

L'Acteur et le système: Les contraintes de l'action collective. By Michel Crozier and Erhard Friedberg. Paris: Editions du Seuil, 1977. Pp. 445. Fr 65.

Arthur L. Stinchcombe
Institute of Applied Social Research, Oslo, and *University of Chicago*

The first task attempted by this book is one that was done for the American tradition of organizational analysis by Everett Hughes, Howard Becker, and Erving Goffman. It takes about 100 pages to say roughly that: (1) personal power arises in organizations only if both the person who exercises power and the person dominated might have chosen other lines of action, and consequently all theories that picture organizational action as determined underemphasize power phenomena; (2) there are nevertheless some very strong normative and realistic constraints on that freedom from which power arises, which tend to appear phenomenologically as "rules of the game," or "situational definitions," or the like; (3) at least on important matters, and to a surprising extent on trivial matters of face and small privilege, people play these games within the constraints fairly rationally, so that once one understands the rules of the game and the more or less standard motives of the people, one can explain most behavior by strategic considerations.

There is a terrible tendency in ethnomethodology and in *L'Acteur et le système* for this approach to lose its lightheartedness and irony and to become ponderously pedantic, with tedious definitions of intuitively obvious concepts. While it would perhaps be harder to do in French than in German or English, one is constantly tempted to tell Crozier that, after all, his maintenance men in the tobacco monopoly were just fooling around. This book refers to the previous work of both authors as what "we" think, have said, etc. A count of the number of times what "we" think comes from Crozier versus Friedberg would be as good an example as the overworked maintenance men, and more amusing.

What strikes me about it, then, is how something I entirely agree with can be made so irritating and pompous that I feel out of character when I believe it. I did not feel out of character when I believed Hughes or Becker or Goffman.

But even with Hughes and Becker and Goffman, there was always the question of whether there was not a bigger game—league standings, perhaps—in which these little games were embedded. The incentive pay game in factories is embedded in the big game of how hard workers will work for capitalist enterprises and how much of the product should go to profit,

how much to maintaining the bourgeois political system, and how much to what the workers themselves choose—see Michael Burawoy's forthcoming work on this general subject, *Making Out on the Shop Floor: Changes in the Labor Process under Monopoly Capitalism* (Chicago: University of Chicago Press, 1979). In Goffman, how much of mental hospitals is a production of total institutions, how much instead is what we would demand of whatever social arrangements we would tolerate to deal with lunatics? In Becker, how much of the authority of principals and teachers is due to the fact that kids really do not know a lot of things educated adults know, and how much is a situational device of the school that could well have been otherwise? Or, more generally, how far are the rules of the little games we play in our roles really explained by those terribly criticizable "functional requisites," which we always argue against except when we are making rules for the institutions where our own values are supposed to be institutionalized. We may be doubtful that societies have functional requisites, but we all make speeches every other week about functional requisites of sociology departments. Whether we speak truly is of no consequence in Becker and Goffman and Crozier and Friedberg. Perhaps our lack of interest in functional requisites is more convincing in this book, because we do not really care whether the monopoly makes enough cigarettes, so we can't get interested in the functional requisites of that task.

But the question we used to pose in our hearts to Goffman ("But what if the patients really are crazy?") has its milder form here ("But how many cigarettes do they really produce?"). The game metaphor starts with the presumption that there is nothing serious going on here, or with the irony that people take winning games more seriously than they do real life. I think we will never understand the other half of the irony of organizational life unless we admit that all those mediocre bureaucrats, on the other side from our guys, are really trying to educate kids and care for the crazy and make cigarettes.

The main point of the second major part of the book seems to be that organizational structures are set up by people and that they set them up the way they are on purpose. That means that any correlation between environment and organizational structure is a choice, a strategy, a human creation. That is, of course, what Philip Selznick and Alfred D. Chandler were trying to teach us, and it is true that the fashion in which Peter Blau, and to a lesser extent Derek Pugh, have tried to make us pretend otherwise in order to be "more sociological" is pernicious. But what this does is to make each of those opaque correlations into a problem: What are people thinking about when they regularly, for example, concentrate authority in the local executive if the local establishment is part of a larger firm, but decentralize authority locally if the local establishment is the whole enterprise? The reason the correlations in the back of Chandler's book are bigger than most of Blau's is that Chandler figured out what to correlate by studying what people thought, what strategies they were following. And I suppose Selznick could have produced correlations with the political

structure of regional authorities if he were not faced with some limitations on the sample size of TVAs. But Crozier and Friedberg do not offer a single explanation of any correlation from the structure-and-environment literature. Theoretically empty statements—for example, that maybe the reason the correlations hold up in several cultures is that they are "not important" or "purely formal" (p. 125)—do not get us anywhere. If people think alike in the same situation in different countries, as shown by these correlations, a book on "the actor and the system" ought to be able to explain it and ought not to wander off into sentimental cultural relativism at precisely the point where the scientific crunch comes.

If I had to guess where the faults of this book come from (the book is by and large right), I would say that they come from the authors' feeling that theoretical originality is a literary effect rather than an illumination of the world. Thus, if we strain long enough to put organizational processes into crisp language, it must be an advance. But without a sure touch, such a way of writing a book degenerates into a manner, and the book comes out too long.

Therefore I will not review the last two-thirds of the book, though I have sort of read it. It is more of the same—good, solid sense stretched out too long. One eventually learns a good deal about French public administration.

Rational Behavior and Bargaining Equilibrium in Games and Social Situations. By John C. Harsanyi. New York: Cambridge University Press, 1977. Pp. x+314. $29.95.

Anatol Rapoport
University of Toronto

A finite game is a model of a decision situation involving two or more actors (players), each of whom must choose one of a finite number of courses of action (strategies) available to him. The simultaneous choices of strategies by all players determine an outcome to which each player assigns a value (utility). A central problem in game theory is that of singling out choices of strategies that can be defended as "individually rational" in the sense of maximizing one's own utility (or expected utility) while taking into account the circumstance that all the other players are also rational, that is, are attempting to do the same. In this way game theory can be regarded as an extension of the normative theory of decision to situations involving conflicts of interest.

Games are classified according to the number of players, according to whether the players can coordinate their strategies to pursue common goals (form coalitions), and in many other ways. Each class of games presents a problem not only of finding optimal strategies (which is a computational problem), but also of defining them in a way to fit the structure of the

game (which is a conceptual problem). For example, in order to define an optimal strategy in a general two-person constant-sum game, the set of available strategies must be conceptually extended to include mixed (probabilistically chosen) strategies. A pair of optimal strategies is said to be in equilibrium: neither player can increase his expected payoff by choosing another while the other chooses his optimal strategy. Such an equilibrium constitutes a determinate solution of the two-person constant-sum game, in the sense of a unique apportionment of expected utility payoffs.

The task that John Harsanyi has set for himself is that of defining a determinate solution for every class of finite games. This task is beset with formidable difficulties, particularly in the context of two-person non-constant-sum games and in noncooperative games with more than two players. Although every such game has been shown to possess equilibria, the problem of singling out a determinate solution among these equilibria remains, as well as the problem of establishing the individual rationality of the choices that intersect in that equilibrium.

The solution of the problem necessitated a modification of the concept of a noncooperative game. The noncooperative game is usually defined as one in which players cannot communicate with each other and hence cannot coordinate their strategy choices to mutual advantage. Harsanyi regards a game as noncooperative if the players cannot make enforceable agreements on coordinated strategy choices. As for communication, it can take place even in noncooperative games in the form of "tacit bargaining," guided by the players' perception of each other as rational actors. This tacit bargaining, based on comparing the risks associated with the failure to agree, leads to a determinate solution. Since in the context of cooperative games determinate solutions reflecting collective rationality (appropriate in such games) have already been proposed (the Shapley Value being an example), Harsanyi's contribution rounds out the formal theory of rational decision in conflict situations.

Harsanyi emphasizes the importance of the rounded-out theory in applications to social science. He writes, "only a theory providing determinate solutions can suggest reasonably specific *empirically testable hypotheses* about social behavior and can attempt to *explain* and to *predict* . . . the outcome of social interaction in various real-life situations" (p. 4).

Here I must disagree. I do not believe that the results of game-theoretic analysis can be used either to explain or to predict behavior in real-life situations. I will go further and maintain that not even the normative aspects of game theory, that is, its prescriptions of how rational actors *should* behave in conflict situations, can be exploited in real-life conflicts. The gap between the game as a model of conflict and real-life conflicts is too enormous to be bridged. It is sufficient to assess what is involved in an empirical determination of cardinal (as opposed to ordinal) utilities of unquantified outcomes to realize that these procedures can be conceived as at best "thought experiments," impossible to carry out with real actors. Over and above these almost insurmountable obstacles is the circumstance that the tacit bargaining that is shown to lead to determinate solutions of

noncooperative games is beyond the capacity of any individual or organization.

The great value of the theory presented in *Rational Behavior and Bargaining Equilibrium in Games and Social Situations* is to be sought not in its descriptive, or its predictive, or its prescriptive aspect, but rather in its potential for enlightenment. Harsanyi builds the concept of rationality from the ground up with admirable clarity and rigor through a large variety of decision situations from those involving one actor to two-person and *n*-person games, cooperative and noncooperative. He treats exhaustively the logical structure of rational bargaining and devotes special attention to paradoxes that lurk everywhere on the path of strategic analysis. Particularly illuminating is his distinction between rational behavior and moral behavior.

In short, Harsanyi brings the full power of mathematical rigor to bear on the analysis of some fundamental concepts of philosophy, psychology, economics, and political science. The mathematical definition leaves no doubt about what one means by a term; the mathematical proof leaves no doubt about the validity of deductions. Thus, two fundamental requirements of a precisely scientific approach to a field of inquiry are fulfilled. The third requirement—empirical validation—is not relevant in the present context for reasons stated above.

A rather high degree of mathematical literacy (but not of technical knowledge) is required to read this book. There is no substitute for this faculty. The social scientist who wants to get some idea about the potential contributions of game theory and related fields to a theory of "rational conflict" will be well advised to learn to read the language, without which the contributions could not have been made, if only to protect himself against much obfuscation and nonsense that has been written and said on the subject.

Sociology and the Demystification of the Modern World. By John Rex. London: Routledge & Kegan Paul, 1974. Pp. xi+271. $14.00.

Dennis H. Wrong
Nuffield College, Oxford, and *New York University*

John Rex thinks that the vocation of the sociologist is, as he insists throughout this book and reiterates in its penultimate sentence, "to demystify the ideology of revolution as well as the ideology of the *status quo*" (p. 253). His argument is generally persuasive, and he is altogether candid and forthright in rejecting alternative conceptions of the sociologist's role, including those of the detached scientist, the policy adviser to princes, the phenomenological anatomist of everyday life, and the *engagé* partisan of radical change. But the term "moral nihilism" that Rex chooses to describe his position is inappropriate. "Moral nihilism"—a pleonasm, surely—is too

strong with its echoes of Nietzsche and the passionate debates of 19th-century Russian intellectuals. Rex is not even recommending the professed value neutrality of scientific positivism, let alone a satanic or Dionysian antinomianism, nor is he bewailing the lost certainties of the past. He wants the sociologist to be ruthless in scrutinizing all moral presuppositions present in sociological work, not in order to extirpate them but because "by destroying, even for a moment, the sacredness of the world view of priests and lay ideologists it opens up the possibility of true moral judgment which is not itself a sociological matter" (p. 9). He does not conceal his own moral stance, which is primarily one of sympathy with the peoples of the underdeveloped continents subjected to economic and military exploitation by both the capitalist and the communist powers.

The "demystification" of the title is also, I think, an unfortunate choice. Weber adapted the term from Schiller to describe the erosion of traditional beliefs and values, especially those of a magical and supernatural character, resulting from the rationalization of modern life, but in the 1960s the word was widely appropriated by radicals and Marxists both inside and outside the academy to legitimate their assaults on the chief agencies of rationalization themselves, to wit, technology and bureaucracy. Indeed, some of them seemed eager to *re*mystify a world drained of magic, mystery, and moral absolutes by the "instrumental rationality" of capitalism and bureaucracy. They were, of course, parroting Lukács and the Frankfurt School, who notoriously cannibalized Weber. Rex sees sociology rather than any class or political movement as the appropriate agent of demystification, but his use of the term is scarcely free from ambiguity. He regards the modern world as the disastrous outcome of the hopes of the Enlightenment, the original source of the Weberian process of demystification. But he also wants to demystify the demystifiers who retain a commitment to revolutionary utopianism. The word cannot bear the weight of these conflicting resonances and serves poorly as a titular description of Rex's project.

The first two sections of the book review the epistemological foundations of sociological analysis and present a substantive theory of social relations, social structure, and conflict that owes much to a careful reading of Weber and elaborates on the themes of Rex's 1961 book, *Key Problems in Sociological Theory* (London: Routledge & Kegan Paul). The third section applies the general concepts Rex has developed to the first, second, and third worlds. The fourth section explores the moral and political implications of the present condition of these worlds and of the sociologist's vocation "in a collapsing civilization."

The two opening sections are a valuable statement of Rex's theoretical position and display a proper skepticism toward current fashions that is rooted in a firm insistence that sociology, if it is to amount to anything, must be substantive and have something to say that illuminates the messy problems of the real world. His argument is dense and terse, insisting, against phenomenologists, that sociologists are not bound to actors'—or "members' "—meanings as the finishing as well as the starting point of

their analysis and that they must develop a conception of social structure that is comparatively grounded. Rex locates the position of quantitative research in this perspective, criticizing the excessive claims made for it without descending to the philistinism of those who chatter about "number-crunching" to justify their own innumeracy. There is a valuable discussion of the variety and versatility of Weber's use of ideal-type concepts and a defense, against Parsonian and Marxist functionalism, of a Weberian view of social structure as a mosaic of "structural bits" that do not necessarily cohere to form a unitary system. All of these themes are discussed at a level that presupposes considerable familiarity with sociological theory, the result being that, like other such efforts, the book fails to achieve the author's expressed hope of making the insights of sociology "available to the great mass of the people" (p. ix).

There is a wide hiatus between the theoretical sections and the review of the contemporary world in the rest of the book. Since the editors of *AJS* asked me to write a review fully four years after the date of publication—presumably because an earlier assigned reviewer was delinquent—it must be said that the book does not wear very well as a "tract for the times." Watergate and Vietnam have passed into history and proved to be rather less typical occurrences than Rex assumed in the early 1970s. Wars and military skirmishes in the third world have broken out, even in a communized Southeast Asia, that are clearly not the results of great power machinations. Black unemployment in the American ghettos has increased, but black violence and crime have diminished, as has white backlash against them. Algeria resembles Frantz Fanon's hopes for it less than ever, and Fanon himself, to whom the book is partially dedicated, now seems to have been the figure of a transitory historical moment, his main heirs the German Baader-Meinhoff and Italian Red Brigade terrorists. The impact of Solzhenitsyn's three-volume study of the Soviet camps has even persuaded some French erstwhile *marxisant* intellectuals to abandon Marxism itself as a totalitarian creed. Several European Communist parties evidently disagree with Rex's evenhandedness in deciding that the differences between the first and the second worlds favor the former despite the "evils" of capitalism.

Yet Rex's pessimism about the future is not necessarily refuted by his overreaction to the events of nearly a decade ago; more recent events, especially in Africa, the continent with which he is most concerned, tend partly to bear it out. But his partisanship on behalf of the nonwhite masses of the southern hemisphere leads him to excessive simplification and selectivity. That the third world is an ideological fantasy rather than a social reality has long been evident. Nor can one push too far Rex's equation of the multinational corporation and the exportable model of the Leninist party as parallel bureaucratic monstrosities. As Rex recognizes on occasion, the first world contains more than capitalism even with respect to its economic institutions and practices. And to identify the second world with the Party as its characteristic institution is to take far too abstract a view, ignoring the actual societies that Communist parties in power have helped

to create, societies from which far more people wish to emigrate or escape to the West, decadent and materialistic though it be, than vice versa.

Sociology and the Demystification of the Modern World has, however, the power of a moving *cri de coeur* combined with the intellectual seriousness of one who does not take lightly the issues posed by his vocation as a sociologist. Possessing a real concern for the peoples of the underdeveloped regions and those of them who have emigrated to Britain, Rex communicates the anguish he feels at being caught between the relative indifference and rapacity of Western elites and the facile attitudinizing of radical sociologists whose commitment both to the cause of the racially oppressed and to a serious view of the vocation of the sociologist is open to question. The grounds for his anguish have hardly disappeared since the early 1970s even if a rather more complex and modulated view of the relations between the West, the Communist nations, and the heterogeneous peoples of Asia, Africa, and Latin America seems indicated.

Ideology and the Ideologists. By Lewis S. Feuer. New York: Harper & Row, 1975. Pp. ix+220. $14.00 (cloth); $3.95 (paper).

Jonathan Rieder
Swarthmore College

Ideology and the Ideologists attests to the constancy of Lewis Feuer's mission of antimythology. Invoked during the End of Ideology scrap, his aphorism declared liberal civilization to start "when the age of ideology is over." He still warns of the dangers to democracy of the dreams of ideologues. A new coinage, "dialectical personality," son of authoritarian personality, places one context of the jeremiad against barbarism in the postwar surge of liberal anticommunism and the image of left and right demonologies fleeing from a vital center to embrace extremism. "For the trails that ideologists have blazed have led to the fires of the Nazi crematoria, and to the Arctic wastes of the Soviet labor camps" (p. 191).

In *The Conflict of Generations* (New York: Basic, 1969), Feuer portrayed his curative sociotherapy "bringing to consciousness the irrationalities" of ideology (p. viii). The title and language suggest the influence of Freud and Feuer's fascination with oedipal jealousies. Yet he has been no remote, inscrutable Freudian figure, more a feisty scold, chastising the Berkeley Free Speech Movement for its parricidal fantasies. He recently savaged hermeneutics as a way of "projecting your ideological myth into texts laden with vacuous words" and borrowed Arthur Koestler's words to attack neo-Marxist dialectics as a "method which enables an idiot to sound extremely clever" (*Contemporary Sociology* 3 [July 1974]: 318). To be sure, Feuer's bite is undialectically clever. Incongruously, *Ideology and the Ideologists* relies on a gauzy Freudian exegesis suspiciously resembling hermeneutics.

Feuer divulges the eternal ingredients in ideology, the emotional a priori. The quest for "invariant structures" owes less to the cognitive structuralism of Lévi-Strauss or to Lacan's Freudian heresy than to a simple notion of the unconscious. Beneath the claims of Marxism, nationalism, and fascism, Feuer detects infantile longing and mythic projection. A biblical allusion expresses the emanationism: "In the beginning was the emotion; then came the idea and last the deed" (p. 79). The sociology of knowledge dissolves in the psychopathology of politics in an erratic, reductive, and engaging analysis.

Ideology is a special vice of intellectuals, a charter for their rule. Weapon of struggle, a cry of pain, mainly a symptom of repressed desire, it is "always an unconscious defence of generational revolt" (p. 82). The wish for social transformation sublimates the rage of the sons for the fathers. Every ideology contains a Mosaic myth. Heroic redeemers—Lenin, Senghor, Mussolini—supplant the elders and establish mystical *communitas*. A father of conservatism, Edmund Burke sensed the theme of castration among the revolutionary "children of their country who are prompt rashly to hack that aged parent in pieces" (p. 91).

Feuer paints the ideologue as a narcissist driven to create the world in his own image. Like Vonnegut's antihero, the ideologue aims to "project a world," but the discipline of science thwarts "the free projection of wishes and emotions" and "evokes a sense of the intrusions of the elders on the 'omnipotence of thought'" (p. 131). The visionary is guilty of bad faith, draping his utopianism in metaphysics, but truth is only decoration. "Virtually every philosophical tenet has been used at some time or other by every ideology" (p. 19). Doctrine flows from a law of generational fashion; the sons of Parsonian progenitors sire conflict theories. The rootless intellectual achieves a conversionary élan by identifying with "chosen people," whether they are blacks, Albanians, or workers. Such passion is disastrous for liberalism. The conviction of election promotes a sadistic propensity to suspend human rights and to sacrifice generations in the name of an exalted humanity. Feuer presents the intriguing thesis that ideology invariably generates anti-Semitism. The purifying impulse compells the ideologue to seek tribal unity, but the Jews resist incorporation into the primal community. They are bearers of an autonomous tradition of individualism and rationality, a liminal race which blurs the boundaries of communal loyalty. They must be annihilated. Jewish survival is thus a metaphor for the promise of civility.

Oddly, Feuer exempts himself from the demands of the balance and evidence he expects from the ideologue. On the strength of a tautology which confines ideology to fanaticism, he removes from analysis the origins of ideology in the suppression of change, conceptions of justice, and social arrangements. Feuer thus repeats the errors of those dogmatic collective-action theories which focus solely on pathology, anomie, and strain. The stress on psyche impedes the comparative study of the role of a variety of factors in shaping ideology, such as party institutions, cultural tradition, and state formation. By the same token, phenomenology and conflict sociology

illuminate a diversity of motives behind generational rivalry which Feuer ignores. The effect is to demote interest and power to the derivative materials out of which myths are fabricated, revolutionary politics to the stage on which psychodramas are ritually enacted. Clifford Geertz summons us back to a pluralistic universe. "Whatever else ideologies may be—projections of unacknowledged fears, disguises for ulterior motives, phatic expressions of group solidarity—they are, most distinctively, maps of problematic social reality and matrices for the creation of collective conscience" ("Ideology as a Cultural System," in *The Interpretation of Cultures* [New York: Basic, 1973], p. 220).

As a counterpart to his conceptual fixity, Feuer violates the premise of empiricism. He resorts to rash generalization and speculative insight. Consider his version of the New Left, which contains a potent dose of truth. A segment of the New Left was afflicted with a nihilistic contempt for liberal civilization, as Feuer charges. The 1967 New Politics Convention endures as a testimony to anti-Semitism, thuggery, and masochism. However, another segment shared the egalitarian ideals of their parents, not oedipal *ressentiment,* and invigorated political conversation by advancing plausible criticism of interest pluralism, corporate power, and American globalism. This segment of the New Left conformed to a democratic heritage which Feuer rightly cherishes.

Feuer's antimythology verges on a mythology of American exceptionalism. He reveres the pragmatism of Jefferson and Franklin and praises America as "a country singularly lacking in ideology" (p. 164) or a myth of historical destiny. This is quite amazing. The "great nation of futurity," America has been influenced recurrently by a millennial idealism and ferocious social movements. The myth of Redeemer Nation has helped to catalyze and rationalize imperial crusades and nativist binges. Our civil religion rests on a Lockean ideology so ingrained that its values of self-regulation and privatism escape public scrutiny. These dilemmas do not detract from the considerable attainments of American democracy. Ultimately, to grant this is to call on Feuer to adhere to his own canons of critical skepticism, to demand a measure of rigor from his passionate humanism, and to affirm the integrity of liberalism.

Social Change in the Twentieth Century. By Daniel Chirot. New York: Harcourt Brace Jovanovich, Inc., 1977. Pp. x+273. $5.95 (paper).

Theda Skocpol
Harvard University

Modernization theory is dead, declared Immanuel Wallerstein at the 1975 Annual Meeting of the American Sociological Association; we should let it "rest in peace" without regret and get on with the business of studying capitalism as a world system. Daniel Chirot certainly agrees, but he does mourn

one thing about the demise of modernization theory: for all of their blatant theoretical inadequacies and ideological misuses, the works of such scholars as Talcott Parsons, Karl Deutsch, and Gabriel Almond did encourage U.S. college students in the 1950s and sixties to take an interest in foreign countries. Since modernization theories have been discredited, Chirot fears, young Americans have tended to turn excessively inward, focusing exclusively on U.S. problems or on issues of personal life-styles. It is past time, Chirot believes, for social scientists interested in comparative macrosociology "to popularize an acceptable and interesting new synthesis" (p. 5). This is needed both to reawaken interest in the study of foreign societies and to encourage educated Americans to realize that "changes in other parts of the world influence American society, just as changes in the United States influence the rest of the world" (p. 6). For these reasons, Daniel Chirot has written a new textbook for comparative macrosociology, one meant "to provide in elementary form a theory of social change that connects changes in the highly developed world with those in less developed societies" (p. 6).

Social Change in the Twentieth Century covers a lot of ground with considerable skill. Arguing that the key fault of modernization theories was to see development and underdevelopment as due to forces purely internal to each society, Chirot instead situates countries in an international hierarchy of core, semiperiphery, and periphery, according to how wealthy, economically diversified, and powerful various countries are in the system as a whole. World capitalism was at its height in about 1900, Chirot argues: capitalist owners (along with privileged middle classes) dominated politics in the core states; core countries had colonized much of the non-Western world; and core investors owned much of the modern, export-oriented economic sectors of peripheral economies. Then the consequences of capitalist expansion generated instability and threats to capitalist hegemony. Economic and military competition among core nations, and between core countries and rising semiperipheral powers like Japan and Russia, led to world wars, depression, and fascism. Meanwhile, core domination unleashed cultural and political forces that led to the revolt of the colonized and semicolonial periphery, generating nationalist revolutions of varying types, which have finally created a phalanx of new semiperipheral states, both communist and authoritarian-noncommunist, determined to assert more independent and active control over the future development of their national economies. For both the earlier and the later phases of the 20th century, Chirot surveys typical stratification profiles of major kinds of countries within the world as a whole. He carefully presents basic facts about patterns of world trade and foreign investments, comparative levels of national socioeconomic development, and statistics on military expenditures. Recurrently, Chirot discusses the prospects of political democracy, its changing meaning, and its suitability (or not) for countries in various international positions. Similarly, a refreshing amount is said about patterns of ideological and cultural stratification and how these have helped to promote changes within nations and on a global scale.

At first glance, *Social Change in the Twentieth Century* seems to offer a "Wallersteinian" analysis. But, in fact, Chirot differs from Wallerstein in crucial ways. Plainly, there are contrasts of political outlook. Whereas Wallerstein seems almost glibly optimistic that the crises of capitalism in the late 20th century presage the coming of a socialist world government, Chirot fears that core capitalist powers, desperate about their waning control over the old periphery and about the rising competition of the communist giants (USSR and China), could evolve toward new forms of fascism. And the only brighter alternative projected by Chirot is for "enlightened isolationism" on the part of a capitalist core newly willing to leave the less developed nations alone to pursue their own economic and political strategies. Associated with such political differences lie important differences in theoretical emphasis. Chirot's approach is more "world historical" than "world systemic"; for Wallerstein, core, semiperiphery, and periphery are enduring structural parts of an all-encompassing world capitalist economy, but Chirot treats these as categories that emerge and shrink over world time. By the mid-20th century, Chirot sees hardly any genuinely peripheral countries; instead, he prefers to speak about the capitalist core and the new semiperiphery, and about the coexistence of world capitalism with communist powers that have closed themselves off, relatively speaking, from capitalist influence (though they have not succeeded in creating an alternative international system). Much more than Wallerstein, Chirot portrays genuine interplay between intrasocietal and intersocietal changes. One gathers from Chirot that changes within countries have done more than allow particular societies to move from one place to another in a fixed international hierarchy; rather, they have "acted back" to change the very structures of international relations. Finally, and perhaps most decisive, Chirot places much more explanatory weight on patterns of political organization (i.e., of national states and revolutionary movements) and on international military competition than would any orthodox Wallersteinian.

Unfortunately, Chirot never raises for explicit discussion his theoretical disagreements with Wallerstein. Indeed, the chief weakness of *Social Change* is its theoretical elusiveness, an ironic fault for a work intended to popularize a new explanation of social change. Partly, the difficulty lies in Chirot's unwillingness to enter existing theoretical debates at a sophisticated level. Not only are differences with Wallerstein glossed over, but Chirot never confronts the arguments of "revisionist" modernization theorists such as Samuel Huntington or S. N. Eisenstadt. Nor does he exploit the theoretical contributions of non-world-systems critics of modernization, such as Alexander Gerschenkron, Reinhard Bendix, and Alejandro Portes. Chirot's own theoretical categories and generalizations cover the sociohistorical realities he narrates only very loosely at best. For example, the British settler colonies (Canada, Australia, New Zealand) are supposed to be "peripheral" (ca. 1900) but exhibit glaringly inappropriate socioeconomic profiles. In addition, much of *Social Change* revolves around accounts of wars, colonization and decolonization, and political revolutions, yet

Chirot never offers any rigorous hypotheses linking these to capitalist economic dynamics. He himself realizes the problem: Leninist hypotheses, derived from the theory of imperialism, are recurrently raised by Chirot as possible explanations for economic-political linkages. But, invariably, Chirot concludes that Leninist arguments do not work objectively (in terms of available data), even though nations "act as if" they believed Lenin's theory were actually true. With this cute paradox, Chirot leaves the theoretical issues unresolved. Knowledgeable readers, if not undergraduate students, will likely conclude that Chirot has dodged the toughest explanatory problems.

In a sense, this is a premature textbook, no matter how badly something of the sort is needed. Textbooks, almost by definition, have a smooth, authoritative air about them. Yet when it comes to rigorously explaining social change in—and of—capitalism on a world scale, all that can be validly done at this point is to raise issues and probe alternative hypotheses. Chirot could have offered significantly more along these lines; he could have produced a more open-ended, intellectually challenging introductory work. Nevertheless, *Social Change in the Twentieth Century* is a very helpful book; used in an explicitly critical and exploratory manner, it will allow teachers to raise the important questions about macrosocial change for their students to ponder with renewed interest.

Interethnic Relations: An Essay in Sociological Theory. By E. K. Francis. New York: Elsevier Scientific Publishing Co., 1976. Pp. xx+432. $22.50.

Jess M. Carrillo
California State University, Los Angeles

Gene N. Levine
University of California, Los Angeles

Reasoning inductively from an explicitly limited range of case materials, E. K. Francis sets out to construct an architectonic theory consisting of an interlocked set of 103 propositions clarified by 59 definitions that would explain "the ethnic aspect in social behavior" (p. 6). This work, abuilding for several decades, is the culmination of the author's career and grows out of his earlier essays, *Ethnos und Demos* (Berlin: Duncker & Humblot, 1965). Francis's central theme "concerns the contradictions between the principles on which the modern nation-state is based and its inability to solve the ethnic problems that these inconsistencies have engendered" (p. xvi). He divides the body of his work into four parts (and 26 chapters): the ethnic factor in social interaction, "ethnie" and polity (wherein the neologism refers to a wider grouping than an ethnic society), interethnic relations in industrial society, and ethnic relations in colonial situations. In part 5 Francis sets forth his "outline of a theory." He aims no less than

"to arrive at a comprehensive set of interlocking empirical generalizations on the basis of evidence supplied by four social sciences [anthropology, history, political science, and sociology], and selected from three continents [central and western Europe, North America, and sub-Saharan Africa] and different historical periods" (p. xii). Francis should not be faulted in this ambitious task for failing to consider ethnic problems and circumstances that beset the other half of the world's populace, that is, experiences in eastern Europe (especially in the USSR and in other socialist countries), in Asia, in the Middle East and the Maghreb, in Australasia and the Pacific, and—save Mexico—in Latin America and the Caribbean. The author builds from a consideration of simpler settings in which the ethnic factor is at play (among the Nuer, the Tswana, the Swazi, the Zulu, and the Ganda), moves on to ethnic groups that cohere in modern (industrial) societies (e.g., the Jews in Minneapolis), goes on to take up the special situations of ethnic groups in nations newly formed out of decolonized cloth (Nigeria, Mexico), and concludes with a consideration of racial segregation and nationalization in the Republic of South Africa.

One of Francis's main distinctions is between primary and secondary ethnic groups. Primary groups focus upon descent and retain cohesion even when their members become subject peoples. His range of cases includes the Manitoba Mennonites, the Mormons, the New Mexico Hispanos, and the South Tyrolese Germans. Signally, the particular situation of the growing Chicano population in the United States finds no place in Francis's work —but then, many a reader will find his pet case absent. Secondary ethnic groups include, in the main, groupings of European migrants to the United States. He considers how well—or poorly—they have been able to maintain their ethnic identity and distinctiveness in the American caldron. European parallels are then drawn (with an emphasis on the Poles in Belgium, in the Ruhr Valley, and in Great Britain) to modify the theorems. Plainly, the author is attempting to specify the sweeping changes members of European ethnic groupings have undergone in the move from tradition-bound to modern, industrial societies. And once the migration has been accomplished, he focuses upon "the structural integration of new additions to a modern society" (p. 295). By integration he means the extent to which members of the migrant ethnic group "function properly within the social order of the host society" (ibid.). Even black Americans are fitted into the category of secondary ethnic subgroups (pp. 283–87). Francis ostensibly uses an equilibrium model in reviewing the fates both of ethnic and racial minorities in Western societies and of decolonized or still subjugated peoples in Africa. He criticizes the "pluralists" (like L. Kuper, M. G. Smith, P. van den Berghe, V. Rubin) who alternatively stress ethnic or racial conflicts that lead to shifts in power (pp. 368–76).

To Francis, the Jews, surely history's classic minority group, present but a "taxonomic problem" (p. 158). Sorting American Jewry by point of embarkation and by recency of arrival, he concludes that while the hordes sprung from the ghettos of Eastern Europe cohere ethnically, those of eman-

cipated, German origin are merely adherents of a religious confession. The point is arguable.

Francis's overall effort at constructing an ethnic taxonomy deserves careful study. Even so, his tightly knit scheme will require modification when conditions outside of his purview are taken into account.

The Ecological Transition: Cultural Anthropology and Human Adaptation. By John W. Bennett. New York: Pergamon Press, 1976. Pp. xi+378. $16.50 (cloth); $9.50 (paper).

Harvey M. Choldin
University of Illinois at Urbana-Champaign

This is an ambitious work of wide scope. It incorporates anthropological materials on small-scale societies, historical studies, and materials about large-scale industrial societies. It draws ideas from biology, systems analysis, anthropology, and other fields, and much of the presentation is at the level of general theory. It is prescriptive as well as theoretical in that the author suggests a path for anthropology to become more useful in the solution of contemporary problems. It appears that the intended audience was professional anthropologists and advanced students, but the book is of sociological interest as well.

Although Bennett calls the work a series of essays, it is a unified volume with a central theme and without internal duplications. It includes chapters on human and cultural ecology, systems and ecosystems, the use of biological concepts, and adaptation as social process.

Bennett offers two principal definitions (p. 156) of cultural ecology, a central concept. The first definition says that cultural ecology is the study of the processes in one major feedback loop in society. The loop involves physical environment, technology, social organization, and "presses," which include values, needs, goals, etc. (fig. 1, p. 38). Bennett emphasizes that the special concerns in ecological analysis are environment, particularly physical environment, and energy-conversion processes. The loop represents "the way the energy-conversion process is related to social and psychological components of the total system" (p. 156).

The policy-oriented definition of cultural ecology states that it is "that field which examines the consequences of human actions toward the physical environment for the environment and for humans, with a view toward modifying or controlling these consequences" (p. 156).

The ecological transition mentioned in the title is "the progressive incorporation of Nature into human frames of purpose and action" (p. 3). It includes the trend toward large-scale societies with population growth and toward mechanization and exploitation of natural resources.

Unlike sociological ecologists, Bennett incorporates values into his ana-

lytical framework as an element of culture. He says that the concept of culture includes values and may include moral precepts "which function as constraints on free choice of models for action" (p. 273). This idea links with his emphasis on adaptation as the criterion for evaluating human ecological arrangements. He presents a master criterion of adaptivity, "sustained yield," which refers to the survival value of the ecological pattern for the collectivity. Beyond this, he offers more specific criteria, such as the biological and psychological well-being of the population as well as the continued good condition of the environment itself (pp. 298–99), recognizing the potential contradictions among the criteria. He encourages anthropologists to apply values and to study the alternative outcomes of ecologically relevant choices.

Bennett is well qualified to attempt to build a cultural ecology, based upon his previous research. He has studied agricultural practices and social change in diverse settings including North America and Japan. The book is grounded in the anthropology of food production and more generally in economic anthropology. The author does not overemphasize his own published work, though, and the volume includes the work of numerous colleagues.

The summaries and critiques of many anthropological works will prove useful to sociologist readers. These summaries deal with two ongoing debates, one on the treatment of environment in anthropological analysis and the other on materialist approaches to social phenomena. Problems with environment as a variable date back to deterministic theories, and Bennett traces the subject through possibilism before reaching his own proposal. The summaries and critiques treat the works of F. Boas, A. L. Kroeber, J. Steward, F. Barth, A. Vayda, M. Harris, and many others.

Bennett's ecology is self-consciously connected with biology. The science of ecology generally emerges from evolutionary theory, which emphasizes adaptation and survival. Even though these concepts should be central to human ecology, they have not been explicitly considered in recent work. Bennett grapples with adaptation and his discussion merits consideration. He also considers the utility of biological concepts in social analysis. He recognizes that bio-ecological ideas have been taken haphazardly from their origins and applied by analogy only. Studies of energy transformation in small-scale societies without high technology represent social scientific analyses which are more consistent with basic ecological ideas.

The Ecological Transition is not flawless, of course. Sociological human ecology gets rather short shrift as Bennett summarizes related disciplines. His conception of sociological ecology is the study of the "geographical distribution of social components in space" (p. 71). This neglects the more general concern with community in sociological ecology. His discussion ignores the line of studies on sustenance organization and division of labor. It also neglects other recent materialist work in sociology.

Bennett omits the built environment as an object of investigation. His scheme shows the physical environment—physiography, climate—as a setting for social life. He neglects the study of vernacular architecture which

shows how people build their own environments, which in turn affect the conduct of social life.

The bibliography is extensive and interesting, containing more than 700 items. There are some errors in it, though: Fred Cottrell appears as Leonard Cottrell and Murray Edelman becomes Maurice. The footnotes are numerous and lengthy but citations in the notes are not cross-referenced in the index.

These are minor complaints about a first-rate book. Human ecologists will find it useful in helping to integrate culture and values into ecological analysis. It also provides an excellent overview of ecological, environment-based, materialist concern in anthropology.

Bureaucracy. By Guy Benveniste. San Francisco: Boyd & Fraser, 1977. Pp. xx+247. $9.95 (cloth); $5.95 (paper).

W. Richard Scott
Stanford University

Guy Benveniste argues that (1) bureaucracies do not work well because (2) they are attempting to relate to increasingly uncertain environments which (3) create excessive fears in their participants causing them to (4) react in a number of defensive and unproductive ways. The first third of the book is devoted to the amplification of these arguments with illustrations drawn from varied types of organizations. A number of games (strategies) played by individual participants and by organizations to insure their own safety and survival are described, as are other games played by client groups in reaction.

The second section is devoted to a description of several current remedies for organizational malfunctioning. The remedies discussed are organizational development (OD), decentralization and increased participation of lower participants in decision making, and the creation of special roles such as ombudsman. Each remedy is judged inadequate to deal with the real sources of the problem. Organizational development approaches deal with individual organizations, not with the environment of organizations. Increased participation is meaningless unless accompanied by genuine decentralization, and this is unlikely to occur to the extent that "centralization emerged because of a need for coordination and integration to reduce environmental uncertainty" (p. 166). And the ombudsman concept can deal with only a limited range of types of organizational errors and, in addition, provides no protection for the organization and its participants.

A final section is frankly speculative and attempts to peer into the future of bureaucracies. Two proposed general ideologies for dealing with organizational issues are reviewed: Large-scale societal planning is contrasted to contemporary self-conscious attempts to reduce the scale of organizations and to simplify social forms. Benveniste notes that planners reduce uncer-

tainty only when they plan correctly and that the movement toward self-sufficiency must either await the development of new technologies or accept a vastly reduced life-style. He sketches out a third schema for reform which involves the creation of a rather complex set of new institutions for oversight and regulation of organizations. Benveniste's admittedly utopian "New System" is composed of boards of professional governors who will protect and evaluate individual careers, councils to protect individual organizations from uncertain environments, professional courts to protect clients and the general public, and professional boards to raise the funds necessary to operate the system and to protect all of us from the new layer of bureaucracies created.

While there are some interesting sections—as, for example, a discussion in chapter eight of the uses and abuses of evaluation research—the argument as a whole is not convincingly made. The key concepts are not clearly defined or consistently used. No evidence is presented documenting the failure of organizations or the extensiveness of fear among organizational participants, or the prevalence of organizational client games. Assertions are made and, often, colorfully illustrated. They are not developed or conditionalized. We are not asked to consider what types of participants or which kinds of organizations employ what types of strategies to survive. In this sense, the great majority of assertions and conclusions proposed in *Bureaucracy* are so vague and general as to be virtually impossible to prove or disprove.

On the positive side, while I found the specific reforms proposed by Benveniste likely to generate more problems than they would solve, the generation of specific alternative futures is an art too little practiced by sociologists. I found the final chapters a refreshing alternative to the conventional sociological denouement: the call for additional research.

The Inheritance of Economic Status. By John A. Brittain. Washington, D.C.: Brookings Institution, 1977. Pp. xiii+185. $9.95 (cloth); $3.95 (paper).

Peter J. Meyer
University of California, Santa Cruz

Let me begin by giving away the ending. John Brittain makes a strong argument that the major determinant of economic inequality is not luck but socioeconomic background (SEB). He also shows that years of education play a strong role, both independently and as an intermediary between SEB and economic status.

Just as he is not a supporter of the "luck" school of inequality, neither is he a meritocrat: That is, he is disturbed by his findings, even if income differences are due to real differences in productivity among people. Nor

is he swayed by arguments that our economy offers individuals equal opportunity to exploit their various productivities.

As long as I am praising negatives, let me add that this book does not address the question of genetic versus parental environmental components of income determination, but not because Brittain shares Bowles and Gintis's view that intellectual skills are not the prime determinant of income. Brittain is not much concerned with the process of education, what is learned in schools, or with socialization processes in school.

This is because Brittain seems to share, in part, the neoclassical economic view of income determination as the result of an endowment of skills and property, although the extent to which incomes are similar to parents' income is due to "the degree of privilege conferred by one's socioeconomic background and other parental influences" (p. 1). Also, his data set does not provide access to answers to questions about nature versus nurture any more than it tells us whether education is "a screening or signaling device to employers" (p. 2) or something that actually changes the students.

This last criticism is only meant to indicate the scope of *The Inheritance of Economic Status:* it is a valuable examination and measurement of the effects of socioeconomic background and parental influence on economic position, not an analysis of the actual processes of upbringing through personnel office procedures years later.

The study is limited in one other way: the sample used. The sample is 659 parents who died in the Cleveland area (Cuyahoga County) and their children, but I would not suggest retitling the book *659 Dead People from Cleveland.* Use of probate records provides hard to come by data while limiting the scope. I have extracted a national sample of parents and children from the University of Michigan Income Dynamics Panel, but it suffers from the young age of the second generation. Brittain argues that his sample is typical, except for religion (majority Catholic), the large size of families, and the origins of the parents (majority immigrant).

On the other hand, the data refer to actual sets of brothers. Brittain makes much of this superiority to Jencks's data, whose central result was a low association between parents' and children's incomes. Jencks did not have actual data on brothers and was forced to infer the correlation between brothers' incomes from equations regressing income on a limited list of background variables. This, of course, biases the Jencks results downward. To the extent that components of background are left out, the effect of background is underestimated.

Brittain is able to compare the variance in incomes of brothers to the variance in the entire sample, partitioning inequality into the part caused by differences in background and the part caused by other effects or luck. Brittain's finding is that family effects account for 34%–61% of the variance in economic status; 34% for a crude measure of occupational status, more for income, and 61% when a composite measure of economic position is used.

Brittain presents essentially the same results in a number of ways. One striking result was that all the sons from the top tenth of the index of

background achieved greater economic status than all the men from the bottom tenth. So much for mobility.

The above results rely on the comparison of brothers to the entire sample and assume that brothers have in common only their background. To the extent that they imitate each other, form partnerships, and so on, the effects are not due strictly to their parents. But then, all these effects are due to the fact that the men are brothers.

Examining particular background variables allows the assessment of the importance of each measurable component of SEB and the expansion of the sample to single-child families. Brittain finds the effects of parental education (and age) to be strongest, although the strength depends on the measure of son's economic position. The effect of parental education is so strong that Brittain suggests it may be a proxy for genetic influences, as well, or for parents' income. My results using explicit measures of parents' incomes and educations together indicate the latter to be true. Brittain also examines parental occupation, family size, religion, and race.

I will deal with the section on the role of education more briefly. By the usual technique of regressing education on background and economic status on education and background, Brittain finds that one-fourth to 40% of the effect of background on economic position is transmitted by means of education (as well as by means of unmeasured variables to the extent that they are correlated with education). Also, he finds a strong independent effect of education.

But is not the effect of education exaggerated because of a more complete specification of education than of background, with which it is correlated? This problem is reduced by the use of the brothers sample, representing background with a dummy for each family. While these dummies do reduce the effect of education (and correlated variables) on income or on a composite index of status, education remains significant. The marginal effect of a college diploma over a high school education was on average a 24% increase in income, holding background constant. I would add that the importance of background to education would make wide differences in education unlikely for people from the same background.

I cannot let pass without comment the interesting half chapter on marital selection, the only section where women are considered. Brittain's finding is that the relationship between a son's economic position and his parents is almost duplicated by the association between his position and his parents-in-law, strong evidence for associative mating. Unfortunately, this section says little about the actual economic situation of women outside the home in the workplace, because the economic status of women is taken to be that of their husbands.

My final reservation concerns Brittain's statements on policy implications. He suggests various tax schemes to reduce the effects of inheritance of both wealth and productivity. I can only say that he seriously underestimates the importance of the reproduction of the distribution of income to maintenance of our economic system.

In summary, I would strongly recommend this book for its empirical

examination of the inheritance of the economic status of men. My disagreements are political; I would go elsewhere for an analysis of the actual processes of reproduction. The most important contribution of the work is to show that inequality is, in fact, inherited.

Industry and Inequality: The Achievement Principle in Work and Social Status. By Claus Offe. Translated by James Wickham. New York: St. Martin's Press, 1976. Pp. 158. $16.95.

Carolyn R. Dexter
Pennsylvania State University

Written originally in German as a Ph.D. dissertation in 1967 and translated for English readers a decade later, *Industry and Inequality* presents an interesting and critical analysis of the validity of the achievement principle as a basis of stratification in modern capitalistic industrial societies. Offe argues that organizational conditions of work have changed in such a way that it is no longer possible to distribute rewards for work on the basis of individual performance. And although much institutional support is given to the achievement principle, it is no longer the basis by which rewards are actually distributed. The concept of achievement has become an ideology for justifying social differences rather than a mechanism for achieving social equality.

Briefly, Offe hypothesizes that modern industrial societies can be characterized by increased numbers of large, specialized, and highly technological organizations and that these characteristics have changed the basis of organizational structures from task-continuous qualifications to task-discontinuous qualifications. By these terms he means that those promoted to the next higher position in the hierarchy require different skills from their subordinates and are often recruited from different sources. This change in organization structure makes subordinates' actions invisible to superiors who are in turn incapable of evaluating performance variations.

These changes in organizational structure modify the workers' role in two dimensions. First, the impossibility of exercising formal interpersonal authority in task-discontinuous organizations requires that organizational control be exercised through internalized norms. Thus the proportion of normative regulative rules increases in proportion to the number of technical rules in workers' roles. Second, specialization and technology have increased the proportion of preventive actions and decreased the proportion of initiatory actions in workers' roles. These changes make performance evaluations impossible as content of the work role has become largely normative. According to Offe, norms are dichotomous. One can only comply or not comply. Moreover they are socially learned and cannot be considered an individual attribute. The individual content of workers' roles has disappeared.

The author validates his argument with detailed analyses of bases of variations in intraorganizational social status: promotion, recruitment, and labor income. In all three cases he finds that peripheral as opposed to central work-role attributes play the predominant role in the assignment of social status. This is especially evident in the case of allocation of labor income, where government intervention in the marketplace through wage actions and secondary programs for income equalization reflects society's shared cultural values more than the actions of individual recipients. Offe feels that relinquishing achievement as a basis of stratification will permit the development of more egalitarian social structures based on status rather than contract—all members of society determining the basis of allocating resources through shared beliefs and values.

Offe supports his arguments with data from both German and American empirical studies. His analysis is wide ranging and detailed and can only be partially presented here. But the arguments are never vague, and his analysis is developed with precision. There is much opportunity for debate on many levels. As he has chosen a central institution, and one seldom studied, as his subject matter, the study is both important and controversial. Scholars working in complex organizations, occupations, professions, and political sociology will all find new insights for their own work as well as new directions for research. The translator deserves special recognition for bringing this monograph to English-speaking audiences.

Comparative Union Democracy: Organisation and Opposition in British and American Unions. By J. David Edelstein and Malcolm Warner. New York: Halsted Press, 1976. Pp. viii+378. $19.95.

Alice H. Cook
Cornell University

Interest in problems of union democracy in this country reached its peak in the early 1960s following the passage of the Landrum-Griffin Act (1959). Edelstein, the senior author of this study and the American member of the team, did his major research at about that time. *Comparative Union Democracy* includes some later data, but fundamentally it rests on material now about 15–20 years old.

Edelstein and Warner introduce their book with a discussion of the nature of oligarchy in trade unions. Their catalog and annotations of American and British writers on this subject are quite complete. Their theory of union democracy follows. Although the subject is necessarily a complex of many variables—they emphasize the hierarchy of officials, regional substructures of the unions, voting systems, and electoral models —they choose to test the existence and degree of democracy in unions by tabulating the number of contested elections for top office and the number of resulting defeats of incumbents. They deal with these data

examination of the inheritance of the economic status of men. My disagreements are political; I would go elsewhere for an analysis of the actual processes of reproduction. The most important contribution of the work is to show that inequality is, in fact, inherited.

Industry and Inequality: The Achievement Principle in Work and Social Status. By Claus Offe. Translated by James Wickham. New York: St. Martin's Press, 1976. Pp. 158. $16.95.

Carolyn R. Dexter
Pennsylvania State University

Written originally in German as a Ph.D. dissertation in 1967 and translated for English readers a decade later, *Industry and Inequality* presents an interesting and critical analysis of the validity of the achievement principle as a basis of stratification in modern capitalistic industrial societies. Offe argues that organizational conditions of work have changed in such a way that it is no longer possible to distribute rewards for work on the basis of individual performance. And although much institutional support is given to the achievement principle, it is no longer the basis by which rewards are actually distributed. The concept of achievement has become an ideology for justifying social differences rather than a mechanism for achieving social equality.

Briefly, Offe hypothesizes that modern industrial societies can be characterized by increased numbers of large, specialized, and highly technological organizations and that these characteristics have changed the basis of organizational structures from task-continuous qualifications to task-discontinuous qualifications. By these terms he means that those promoted to the next higher position in the hierarchy require different skills from their subordinates and are often recruited from different sources. This change in organization structure makes subordinates' actions invisible to superiors who are in turn incapable of evaluating performance variations.

These changes in organizational structure modify the workers' role in two dimensions. First, the impossibility of exercising formal interpersonal authority in task-discontinuous organizations requires that organizational control be exercised through internalized norms. Thus the proportion of normative regulative rules increases in proportion to the number of technical rules in workers' roles. Second, specialization and technology have increased the proportion of preventive actions and decreased the proportion of initiatory actions in workers' roles. These changes make performance evaluations impossible as content of the work role has become largely normative. According to Offe, norms are dichotomous. One can only comply or not comply. Moreover they are socially learned and cannot be considered an individual attribute. The individual content of workers' roles has disappeared.

The author validates his argument with detailed analyses of bases of variations in intraorganizational social status: promotion, recruitment, and labor income. In all three cases he finds that peripheral as opposed to central work-role attributes play the predominant role in the assignment of social status. This is especially evident in the case of allocation of labor income, where government intervention in the marketplace through wage actions and secondary programs for income equalization reflects society's shared cultural values more than the actions of individual recipients. Offe feels that relinquishing achievement as a basis of stratification will permit the development of more egalitarian social structures based on status rather than contract—all members of society determining the basis of allocating resources through shared beliefs and values.

Offe supports his arguments with data from both German and American empirical studies. His analysis is wide ranging and detailed and can only be partially presented here. But the arguments are never vague, and his analysis is developed with precision. There is much opportunity for debate on many levels. As he has chosen a central institution, and one seldom studied, as his subject matter, the study is both important and controversial. Scholars working in complex organizations, occupations, professions, and political sociology will all find new insights for their own work as well as new directions for research. The translator deserves special recognition for bringing this monograph to English-speaking audiences.

Comparative Union Democracy: Organisation and Opposition in British and American Unions. By J. David Edelstein and Malcolm Warner. New York: Halsted Press, 1976. Pp. viii+378. $19.95.

Alice H. Cook
Cornell University

Interest in problems of union democracy in this country reached its peak in the early 1960s following the passage of the Landrum-Griffin Act (1959). Edelstein, the senior author of this study and the American member of the team, did his major research at about that time. *Comparative Union Democracy* includes some later data, but fundamentally it rests on material now about 15–20 years old.

Edelstein and Warner introduce their book with a discussion of the nature of oligarchy in trade unions. Their catalog and annotations of American and British writers on this subject are quite complete. Their theory of union democracy follows. Although the subject is necessarily a complex of many variables—they emphasize the hierarchy of officials, regional substructures of the unions, voting systems, and electoral models —they choose to test the existence and degree of democracy in unions by tabulating the number of contested elections for top office and the number of resulting defeats of incumbents. They deal with these data

in two ways: statistically, and descriptively in case studies. The first approach results in a data study of 31 British and 51 American unions of given sizes as of 1952 and 1954, respectively. (They offer no explanation of why they chose a statistical base more than 20 years old, even though their case material covers the sixties and makes passing reference to even more recent events. A possible explanation is that they wanted to use—or had access to—the same data which the British scholar, V. L. Allen, had analyzed in 1954, but the study is not designed to replicate Allen or to compare the respective conclusions.)

Before undertaking the statistical analysis, Edelstein and Warner somewhat laboriously develop a set of hypotheses. Their tests of these produce no startling results. In Britain there are more candidates for top office than in America, more decentralization in union structure and power, less overt organization of factions, more use of proportional representation in voting, and more runoff elections. But are the British unions therefore more democratic? The authors find that "in view of the overwhelming weight of factors favourable to electoral opposition in British unions, it is puzzling that the British periodic elections are not much closer than the American, and that top officers are more rarely defeated in Britain than in the United States" (p. 133).

The case studies that make up the second part of the book are its better half. The two British chapters are classically good. They do not limit themselves to elections but reveal many elements of union life not anticipated in the hypotheses. A distinctive and recurring factor which seems decisive in Britain is the influence of outside political groups—Communist, Maoist, Labor, left-socialist—with which contestants are identified and from which they draw unmeasured amounts of aid and support. The unanswered questions are how much do these political groups account for a multiplicity of candidates and to what degree do they contribute to union democracy?

The case study of the British engineering union's elections shows that, while they have been actively contested and have often resulted in runoffs, they have engaged the participation of a very low percentage of the members—7.9% in two of the elections and not above 15% in 30 years. Participation may not determine the outcome of elections, but surely it is at least as good a measure of democracy as is the fact that elections are contested. The significance of participation is not hypothesized. One learns, however, that Edelstein and Warner prefer the referendum to balloting in convention and the postal ballot to casting votes at the union hall or workplace. To this extent participation has its place.

A factor favoring high rates of contest in Britain is determined to be the existence of fairly autonomous subdistricts within the union, where regional leaders develop constituencies and gain political experience. Unfortunately, the essay on American cases does not examine this factor, and we are thus deprived of an opportunity to understand the significance of this structural element in a comparison of the two systems.

The book as a while is diffuse. Its two parts are insufficiently connected and interwoven; and its concluding chapter, "The Future of Union De-

mocracy," moves far afield from the thesis in order to deal with such matters as humanization of work, worker-management, rationalization, and even to speculate on the role of unions under a possible democratic socialism. More serious is the questionable value of its basic thesis as a decisive indicator of democracy. On the positive side are the two British union case studies, the summary of writings on union democracy and oligarchy, an excellent bibliography, and a discussion of the differences in union structure and function between the United States and Great Britain.

Modeling Social Processes. By Patrick Doreian and Norman P. Hummon. New York: Elsevier Scientific Publishing Co., 1976. Pp. xvi+172. $10.50.

Robert L. Kaufman
University of Wisconsin–Madison

In *Modeling Social Processes*, Patrick Doreian and Norman Hummon argue for the use of a systems-theoretic approach for the study of social processes which evolve over time or space. They distinguish among three components which together define a system. The units (or variables) are the individual pieces of which the system is composed. The structure refers to the relationships among the units of the system (including the distribution of the individual units), while the process is "the means whereby structures are changed or maintained" (p. 6). In this perspective special attention is paid to change and to the mechanisms which control change as the focal points for theory and research.

Using this systems approach, Doreian and Hummon provide a coherent rationale for the use of systems of differential equations to model the through-time (or space) trajectories of the units of the system. Systems of differential equations are ideal for many applications since: (1) They distinguish between units, structure, and process; (2) it is possible to incorporate information on the environment of the social system into the model; (3) once the parameters of the model are known, the dynamic properties of the system (projection over time, equilibrium/disequilibrium) can be studied; (4) the effects of a change in any variable in the system or in the environment can be predicted for the total system.

Doreian and Hummon present four examples of differential equation models. For each example they outline the theory underlying the particular model in conjunction with the mathematical specification of the differential equations. They then develop estimation procedures and discuss the properties of the model as given by the parameter estimates. Three of the models (labor turnover, organizational change, and government change) explicitly allow change only over time, while their model of political insurgency incorporates change over both time and physical space. However, all of the models are based on the authors' notion of "structural control." The basic idea behind this concept is that for each variable in

the system there is an ideal value at any point in time which is determined by the observed values of the other variables in the system at that time and by the values of the exogenous (environmental) variables. Change in the system, then, is generated by the discrepancies between the ideal and the observed values of each variable endogenous to the system.

There are, however, three important issues which the authors do not address; two concern their use of differential equation models and one concerns their estimation procedures. First, there is the issue of the appropriateness of continuous-time as opposed to discrete-time models, that is, of differential equation models versus difference equation models. For their example of labor turnover in particular, it is difficult to conceive of their endogenous variables (labor turnover and overtime) as changing instantaneously. Yet using differential equations to model changes in these variables explicitly assumes instantaneous change. Instead, it seems more appropriate to conceive of these variables as changing periodically (e.g., day to day) as step functions over time and to model the system using difference equations. System of difference equations have the same desirable properties as mentioned above for differential equations and are clearly appropriate where change can occur only periodically and not continuously (see also W. J. Baumol, *Economic Dynamics* [New York: Macmillan, 1970]; S. Goldberg, *Introduction to Difference Equations* [New York: Wiley, 1958]).

Second, the models which Doreian and Hummon develop based on their concept of "structural control" are mathematically indistinguishable from an alternate model. Suppose that a model is developed, on theoretical grounds, that states that the change over time in an endogenous variable, say, the first [$dY_1(t)/dt$], is a function of the observed value at time t of that endogenous variable [$Y_1(t)$] as well as of the observed values of the other $k - 1$ endogenous variables [$Y_2(t), \ldots, Y_k(t)$] and the values of the exogenous variables ($X_1, \ldots, X_n$). That is,

$$dY_1(t)/dt = a_1Y_1(t) + a_2Y_2(t) + \ldots + a_kY_k(t) + b_1X_1 + \ldots + b_nX_n \,. \quad (1)$$

Using equations analogous to (1) for the other $k - 1$ endogenous variables, a system of differential equations can be specified which is mathematically identical to the kind of models the authors develop. But the interpretation of the two models is quite different. For example, under the "structural control" model the parameter a_1 indicates the sensitivity for change in $Y_1(t)$ to discrepancies between the observed and ideal values of $Y_1(t)$, while under the alternate model, it describes a feedback mechanism for changes in Y_1 being generated by the observed value of Y_1 at time t. Thus, their "structural control" models cannot be tested against an alternate hypothesis of feedback as described above.

Third, there is an issue concerning their estimation procedures which poses a severe problem for their parameter estimates. While they do consider and resolve the problems which can occur if there is autocorrelation among the errors for a single endogenous variable, they do not consider the

problem of between-equation error covariance. Among other reasons, it is eminently likely that there will be between-equation error covariance, as this would capture effects on the endogenous variables which are unique for each observational unit. Then, given that each equation contains lagged endogenous variables which will be correlated with the error term and given the between-equation error covariance, an OLS or even a GLS estimation procedure for any single equation (which the authors employ for all their examples) will produce parameter estimates which are neither unbiased nor even consistent (see also H. Theil, *Principles of Econometrics* [New York: Wiley, 1971], chaps. 8–10). Thus, their estimation procedures produce parameter estimates which must be viewed with skepticism as to their accuracy.

In sum, Doreian and Hummon provide a good introduction to the use of differential equation models for social systems, but their work suffers from the neglect of several important issues.

Women in the Twentieth Century World. By Elise Boulding. New York: Halsted Press, 1977. Pp. 264. $15.95.

Carolyn M. Elliott
Wellesley College

"Running out of women" would be more serious than running out of oil, according to Elise Boulding's assessment of women's contributions to food systems. Women form a vast army of growers, transporters, sellers, processors, and servers of food. In Africa, up to 70% of agricultural production (p. 114) is done by women; elsewhere women's labor is the major transport system for water, the primary component of any food system. Both equity and progress demand that these socially invisible roles of women be recognized. Boulding's image of a nursing mother serving different meals to infant, small child, husband, and often festival guests while denying herself adequate nutrition is a powerful and poignant one. To those who might argue that progress demands such sacrifice or, apparently more benignly, that development cannot proceed on all fronts at once, Boulding makes a more difficult and significant argument. Because of their culturally different experience, women have special knowledge of local social terrains and the time and energy demands of innovation that planners cannot afford to ignore. Women also know, unlike socialist planners, that "the private shapes of individual lives cannot be fully mass-serviced" (p. 71). Women's resourcefulness and ability to meet individual needs from local materials are needed not only by developing countries. The world system as a whole needs to learn how to replace poorly functioning centralized systems with self-help techniques practiced by women householders.

Boulding's argument is presented in a series of 10 essays, six of which were issued previously and rewritten for *Women in the Twentieth Century*

World. The essays are part of a massive many-faceted project which has recently resulted in both the *Handbook of International Data on Women* (New York: Wiley, 1977), coauthored with Shirley Nuss, Dorothy Carson, and Michael Greenstein, and her own macrohistory of women, *The Underside of History* (Boulder, Colo.: Westview, 1977). Among the essays are Boulding's widely quoted "Women, Bread and Babies," as well as analyses of economic dualism, food systems, nomadism and urbanization, and familial constraints on women's work roles. The final section of the book changes from these socioeconomic topics to a consideration of women's international voluntary organizations (NGOs) and their capacity for acting as world peacemakers.

The reach of these essays is truly macrohistorical and global. Boulding ranges from prehistoric agro-villages and present-day nomadism through the impact of urban civilization to futurist planning in a world society. In the course of these discussions, she provides many arresting concrete insights which locate her analysis in the particularities of daily life, a style which well matches her argument. She alerts us, for instance, to the nurturant roles played by men in Laslett's 16th-century England, where families raised each other's children as craft apprentices, and to the communications function of hairdressers working among women. She further helps us to recognize Dickens's ragged children as the children of working mothers in a society which made no accommodation for women bearing family responsibilities alone.

The methodological orientation of the book is toward analysis of aggregate data, drawn largely from the *Handbook* of UN census data by Boulding et al. Boulding's two major interests are measuring the work done by women, both in the paid labor force and in the subsistence sector, and assessing women's contribution to the household. The difficulties of doing so are as prodigious as the urgency, both of which Boulding amply demonstrates. I find the results of her efforts, however, disappointing. She often presents too much information, overwhelming the data needed to support her analysis with statistics which prove to have no correlation. There are other occasional problems of presentation, including noncomparable tables (pp. 122, 126), misquoting data from tables (p. 123), and a confusing method of drawing data from more than one place to make a single point (pp. 122–24). The book would be easier to read if the tables had been redrawn for the purposes of this presentation and the supplementary data placed in appendices.

A more important concern, however, is the reliance on aggregate data. Boulding's analysis shows the insufficiency of aggregate data analysis, for her statistics often fail to illuminate her questions. This is most apparent in her discussion of the impact of economic dualism on women, in which she relies on the controversial indices of dualism developed by Adelman and Morris. After an extended discussion, Boulding admits that she finds no real relation in the data between dualism and women's labor force participation, because women show both high and low participation at every level of dualism. One senses with Boulding that there probably is a very

significant relationship somewhere. To find it, however, we must turn to detailed studies of relationships between classes and occupational groups within countries rather than summary indicators of whole societies.

Boulding's essays are significant because they help us to see things differently. She provides many arresting metaphors—for example, that working women are moonlighting when they return home to process food—which give concreteness to her very important analysis of women's contribution to food systems. The next step is an examination of the concrete processes that make up these systems in particular contexts in order to interpret the meaning of the statistics she reports.

Power and Pawn: The Female in Iberian Families, Societies, and Cultures. By Ann M. Pescatello. Westport, Conn.: Greenwood Press, 1976. Pp. xix+281. $15.95.

Terry J. Rosenberg
State University of New York at Stony Brook

Ann Pescatello has, by her own admission, chosen a task ". . . presumptuous in its temporal or its spatial dimensions" (p. xv): a survey of female roles in Iberian societies, past and present. In spite of its seeming unmanageability, she persisted, spending over a decade on the data collection and final report. Sadly, Pescatello's efforts offer limited new information and less new insight. Her study is disappointingly superficial. It is atheoretical, methodologically unsophisticated, and marred by careless editing and proofreading.

A review of the chapter headings immediately impresses the reader with the monumental scale of Pescatello's endeavor. In the first chapters, she deals with women's roles in Spain and Portugal before the conquest of the New World. Next she describes the social, economic, political, and religious contexts of the areas which underwent Iberian exploration—Asia, Africa, and Amerindia. The confrontation of native and European heritages during the conquest is briefly discussed; continued developments through the 19th and 20th centuries in all societies influenced by Iberian settlers are then summarized. Finally, one chapter is devoted to an examination of women's roles in 20th-century Brazil, before the book closes with an epilogue "in the manner" of a summary and conclusion.

Pescatello's sources appear in a massive bibliography, which includes works in at least half a dozen languages. Many of these are primary sources, predominantly reports of early travelers and colonizers. Others are secondary historical analyses. While some data are taken directly from censuses and government surveys, more often statistics are derived from official or scholarly summaries. Conspicuously lacking in the bibliography are recent books and articles (in any language) by other social scientists. The bulk of recent references consists of articles published in an earlier collection Pescatello edited. Although years of field experience are cited

as another source, the text contains meager evidence of her personal observations.

As a whole, the book is overburdened with broad generalizations and minute statistics which are never integrated into a meaningful argument. Pescatello informs us that prior to conquest "Africans had developed sophisticated social, political, economic and belief systems" (p. 87). We are told that "Uruguayan boys and girls apparently are taught sex roles at an early age . . . " (p. 203). In the way of unnecessary detail, we find a full listing of 1914 female employment figures for Argentinean cities, with no obvious reason for considering this dated regional breakdown.

Pescatello could have given us well-prepared, useful, summary tables of the pertinent information which is already present, but not easily accessible, in her text. For example, I would have appreciated tables showing national statistics on female employment, dates for the passage of important national legislation affecting women (e.g., the franchise, protective labor laws), and charts of the relative rates of progress through educational institutions for women and men in the countries studied. These data are now scattered through the text like pre-Columbian artifacts on an Andean mountainside; the reader should have been spared an archaeological trek in favor of a comfortable museum visit.

Interesting anecdotes and historical commentary dot the first few chapters. The traditional seclusion of upper-class Portuguese women is charmingly revealed in an old proverb: " . . . a virtuous woman left her home only three times during her lifetime: for her Christening, her marriage, and her funeral" (p. 157). Likewise the odd perception of Iberian colonizers is highlighted by the words of a Spanish bishop who, in 1566, marveled at the large breasts of Mayan women and attributed their special endowment to " . . . the continued grinding of tortillas [which] agitated their breasts . . . " (p. 124). The intriguing subjects of female slavery, female prostitution, and the growth of Catholic convents are treated in the first chapters. Pescatello portrays slavery and prostitution in both the indigenous and the Iberian cultures and then follows their natural emergence in the hybrid colonial setting. She explains how the confinement of nuns was often used as a form of population balance in societies with an excess of females. In other societies, she describes the competition between the Church and eligible bachelors over the few available European women with substantial dowries. These observations occasionally refresh the weary reader.

The most disappointing chapters of the book are the ones dealing with the contemporary period. It is disheartening to see Pescatello rely so heavily on area handbooks for information. The statistics from these and other secondary sources are inexcusably outdated. Here the need for current references from economics, sociology, anthropology, and political science is most obvious. Two works, among many others, whose absence is startling are E. Boserup's *Women's Role in Economic Development* (New York: St. Martin's, 1974) and N. H. Youssef's *Women and Work in Developing Societies* (Westport, Conn.: Greenwood, 1976).

In the contemporary chapters, Pescatello also fails to mention a few noteworthy issues. She hints that in Argentina and Chile there is some divergence from the strictly Iberian pattern of sex roles, but she overlooks the contribution of recent European immigrants to these national cultures. Even more glaring is her neglect of the revolution in fertility and family planning which has been fully documented for many countries of Central and South America.

Though Pescatello claims that " . . . in the figure of the mother and thus as 'head' of the household [the Iberian female] has been the central figure in the family and thus in society" (p. 231), there is scant information presented to support her contention that women are power brokers. Precisely to the contrary, she repeatedly offers evidence of the sexual and emotional imprisonment of women, of the exploitation of women as low-level workers, and of the legal limitations on the civil rights of women. She shows how remarkably persistent the legal definition of women's inferiority has been in Iberian societies. From the days of the conquest, she notes: "Iberian law codes identified women as 'imbecilitas sexus,' an imbecile by her very nature . . . women were classed with children, invalids and delinquents" (p. 141). This inferior status survives today, for example, in the Brazilian labor code which provides that " . . . husbands and fathers can bring action to terminate the employment of wives and daughters under 21 on the grounds of family disruption" (p. 219)

The "powers" enjoyed by women which Pescatello could have more fully elaborated include the widow's control of her deceased husband's property and the management of her impoverished household by the lower-class woman. We are briefly told of unusually powerful widows in indigenous tribes and in colonial families who assumed all of their husband's perquisites. We are also told that lower-class women have had more mobility outside the home, more personal freedom, and a kind of power their elite sisters were always denied because of the formers' economic responsibilities and the instability of their family life. But, at the same time, Pescatello concludes that "women who belong to groups on the periphery or totally outside the ken of power are as powerless as the groups to which they belong" (p. 232).

It is strange that Pescatello ignores the potential power of women within the economic arena. While she cites the nearly universal status of Iberian women as agricultural workers, domestic servants, and industrial laborers, she does not acknowledge either the centrality of this role or its importance for the whole economy. Mention is made of the black heritage of female participation: "[West African] female control over marketplace production and distribution has been a central feature of Afro-Latin, particularly Afro-Brazilian, cultures" (p. 89). For other areas of South America similar early involvement is noted: "In Colombia's Cauca Valley women dominated weaving enterprises and also totally controlled a labor pool of children under twelve . . . " (p. 122). In Argentina, Brazil, and other countries women and children are seen as the mainstays of nascent textile and food production industries. Unfortunately, though, Pescatello never focuses on

Iberian women as workers, a role which might have enhanced their power.

Pescatello proposes a new history. "We must," she says, "move away from the study of the narrow male elite groups and shift our attention to the heretofore inarticulate, the power *base* of history as it were rather than the power *manipulators*" (p. 233). She has not, in my opinion, advanced the cause of this new history with *Power and Pawn.* Certainly, she and other historians are handicapped by the very inarticulateness of the power base: if women have been illiterate and uneducated for centuries, they have been unable to transmit their own interpretation of historical events through normal channels. Even with such deficiencies, however, pieces of the historical picture that Pescatello tries to paint with a barn-sized roller can probably be studied in much greater depth than she has accomplished. Within the realm of social science, as a serious scholarly treatise, Pescatello's broad approach seems premature.

Pink Collar Workers: Inside the World of Women's Work. By Louise Kapp Howe. New York: G. P. Putnam's Sons, 1977. Pp. 301. $8.95.

Sharon N. Barnartt
University of Texas at San Antonio

Ingrid, a waitress, is overworked and underpaid. Jackie and Linda, beauticians, and Peggy, a sales clerk in a large department store, are underpaid. Joyce and Bea, homemakers, are overworked but unpaid. Bonnie, an office worker, is underpaid—but her company provides free lunches. All of these women have male bosses, no unions (or else male-dominated unions), and few job benefits. With the exception of the homemakers, their work is so rigidly regulated they have little chance to express whatever creativity their tasks could allow. Their work, however, is not unskilled or undemanding. In this book, these women and others tell us about the ways occupational segregation affects their work lives. Regardless of their jobs, they have common complaints: economic exploitation, job insecurity, boredom, powerlessness, the impossibility of advancement, and the inhumanity of many of the (usually male) supervisors.

Pink Collar Workers is about some of the jobs traditionally filled by women. (Blue-collar jobs are filled primarily by men, pink-collar jobs by women.) Louise Kapp Howe focuses on only those jobs usually held by women without college degrees. These jobs are filled by approximately 60% of the paid female labor force, according to the Bureau of Labor Statistics (Bulletin 1880), and by the 60% of women who are not part of the paid labor force (i.e., homemakers). She does not consider women who work in "women's professions" of nursing, teaching, and social work or those who hold traditionally male jobs.

Howe does not attempt a thorough sociological analysis of occupational

segregation. Rather, she describes the results it has for women who work in pink-collar jobs through (presumably) verbatim interviews with women who work in five pink-collar occupations. These women articulate the problems and the possibilities of their jobs well, and because of them this is a very powerful book. Were it about past events rather than present realities, it would be called oral history at its best. In addition, the book is not devoid of sociological analysis. In her introductory and concluding chapters, through judicious editing, and in a running commentary, Howe explains why these occupations are structured as they are, who profits, and how the women lose. This analysis shows the ways in which the structures of the occupations are oppressive as well as the ways in which women contribute to (or at least accept) their own oppression. She treads delicately between blaming the men in power and blaming the system itself, but she does not adequately explain why the women do not, cannot, and, until very recently, have not attempted to change things. Her informants do, however, perhaps unwittingly, tell us why such attempts will meet resistance from women themselves as well as from those in power. As such, the book should make committed feminists despair. While it shows vividly why changes are needed, it also demonstrates the power that socialization and social institutions have in causing women to accept their own exploitation.

The major flaw in the book is its anti-intellectual bias. Howe has addressed the book to women without college degrees who are the most likely to work in pink-collar jobs, and she clearly feels that these women suffer more than women in other female jobs or in traditionally male jobs. She does not consider the possibility that women in all jobs share many of the problems of women in pink-collar jobs, whether they are in female or male occupations. But this bias also prevents her from identifying what is perhaps the most exploitive result of occupational segregation for women, and that is the degree to which women's educations are underutilized. Howe ignores the dual facts that education does less to improve wages or working conditions for women than it does for men and that many of the women she interviewed were capable of doing work which could have been more satisfying as well as more rewarding.

Lack of attention to this aspect of occupational segregation makes both the description and the analysis incomplete. However, the book cannot be ignored, because it fills an important gap in the literature on women and work. Lower-status women's occupations have been ignored by sociologists even more than lower-status men's occupations have. This book paints a vivid portrait of women's work life in pink-collar occupations and must be part of any consideration of women in the labor force or of the sociology of work.

Sex Roles: Biological, Psychological and Social Foundations. By Shirley Weitz. New York: Oxford University Press, 1977. Pp. xi+283. $10.00 (cloth); $5.00 (paper).

Inge Baer Corless
Russell Sage College

Shirley Weitz has written a useful text for undergraduates studying sex roles. Her particular emphasis is on examining how a variety of biological, psychological, and social processes tend to maintain the "sex role system." Sociologists may find her discussion of the biological and psychological considerations interesting and informative but are likely to deem her treatment of social and cultural questions inadequate.

Weitz organizes her analysis by examining what she terms "The Biological Maintenance System"; "The Psychological Maintenance System"; "The Social Maintenance System: The Family"; "The Social Maintenance System: Symbolism"; and "Sex Role Change through Space and Time." For each of these topics, Weitz considers some of the key issues. For example, in considering the Biological Maintenance System she examines aggression, sexuality, and psychosexual abnormalities for evidence pertinent to sex-role differences. She is more successful when she uses this approach in the first two sections of the book, where she provides a neat, competent review of the relevant findings in biology and psychology. An omission of some importance, however, is the failure to address the work of the sociobiologists. The latter three chapters, on social and cultural considerations, do not come up to the same high standards of the first chapters.

A weakness in her conceptual approach is that, although the discussion in this text is organized around maintenance systems, these systems are never defined and seem to be used merely as chapter headings. The term "system" is also used with reference to biology, psychology, and sociology, as well as for sex roles. It would be helpful analytically if these concepts were carefully delineated. Particularly with reference to maintenance systems, the discussion might profit by demonstrating its theoretical linkages to Parsons's hierarchy of systems. Such an exploration could provide a powerful vehicle for discussing the interpenetrations of these maintenance systems.

Another serious problem is the limited questioning of the status quo. This is particularly puzzling when Weitz is concerned with producing change (p. 5). For example, why is the problem of aggressiveness not considered to be a social problem rather than merely a problem in males? And why is the menstrual cycle, which is perceived by some as disqualifying a woman for high public office, not contrasted with aggressiveness, which would seem a potentially more serious disability?

On another occasion Weitz states that "the differing consequences of sexual activity for men and women, short-term pleasure for one, possible pregnancy and long-term maternal responsibility for the other, have doubt-

less largely determined the nature of societal restrictions and structures" (p. 29). Maternal responsibility determines societal restrictions for Weitz. She does not consider that societal restrictions may determine maternal responsibilities. Evidence for such an interpretation is seen in the continual lengthening of the dependent period of children in our society. Moreover, sexual activity for women includes not only possible pregnancy but also pleasure and the power of giving life. In addition, monogamy and more specifically patriarchy might be conceived as capturing the power of life giving, that is, males possess the power by possessing the females.

A further illustration of this emphasis on the status quo is given when Weitz writes that "traditional attitudes linking men with the public sphere and women with the private, domestic one seem to underlie much of the resistance to sex role change" (p. 230). There is no attempt to explain the reasons for the maintenance of these attitudes, such as their service in restricting the competition for scarce resources in education and in professional careers.

Continuing with a consideration of Weitz's approach in examining the critical issues, the chapter on symbolism would benefit from additional attention to the arts, namely, literature, film, and theater. Television and advertising are other areas reflecting and reinforcing sex roles and thus also require investigation. The last chapter, entitled "Sex Role Change through Time and Space," takes us backward but not forward. For some reason Weitz avoids giving us a reason for the persistence of sex-role differences or an indication of what is probable in the future. A petty annoyance which is indicative of her lack of sensitivity to the use of sociological materials is Weitz's saying that she will examine sociological data and then introducing an anthropological study. On other occasions, she mentions a sociological system when what she may mean is a social system.

Minor disagreements with Weitz in terms of interpretation include her description of Jan Morris as having a "distaste for masculine activities" (p. 52). Au contraire! She seems to idolize men and their activities and to pity women. Also, with reference to Louisa May Alcott's *Little Women*, Weitz cites this classic as reaffirming "values of feminine domesticity" (p. 88). Perhaps, but we were also presented with a tomboy role model in the form of Jo, an individual who had intellectual aspirations.

Despite these limitations, *Sex Roles* may prove to be a valuable text for use in undergraduate courses. Most of the biological and psychological issues are discussed thoroughly, and, even though the treatment of social and cultural factors is more superficial, the material is likely to raise engaging and stimulating questions in the context of class discussion—which may have been Weitz's intent. We still need such texts in the area of sex roles.

The Black Male in America: Perspectives on His Status in Contemporary Society. Edited by Doris Y. Wilkinson and Ronald L. Taylor. Chicago: Nelson-Hall, 1977. Pp. viii+375. $17.50 (cloth); $6.95 (paper).

Walter R. Allen
University of North Carolina at Chapel Hill

Wilkinson and Taylor sought to "focus upon the social positions and role enactments of African American males" (p. viii) in contemporary American society. What results is a publication of considerable significance and importance to scholars concerned with the study of the black experience in this country. Drawing upon the work of sociologists, social psychologists, psychiatrists, and anthropologists, the editors manage to examine from an interdisciplinary perspective many of the myths and realities surrounding black men. Among the issues addressed in this unique volume are: the roles of black males as fathers, husbands, and providers in their families; male socialization processes in black communities; the form, functioning, and consequences of female-headed households for black males of all ages; the nature of societal barriers which restrict full economic participation by black males; black male–black female interpersonal relationships; interracial mating by black males and sex-role orientations in the black community. The book is organized into four sections: (1) "Socialization to the Black Male Role," (2) "Stigmatization: Systematic Branding," (3) "The Issue of Interracial Mating," and (4) "Roles and Statuses in Post-Industrial Society." Underlying and integrating these sections is a common concern with the social statuses and roles assigned to black males by this society. Through a masterful blending of original papers and selected works that have appeared elsewhere, Wilkinson and Taylor bring to bear on the black male condition in this country an enlightened, perceptive analysis quite unlike any I have seen before. When one considers the paucity of quality research which has been done on black males to date, this is no small accomplishment.

In order to convey better the substance of this timely book, I turn now to a brief summarization of its contents. Section 1 begins its consideration of black male socialization experiences with David Schulz's article, "Coming Up as a Boy in the Ghetto." Schultz's conclusion (reflective of what Ronald Taylor calls the "orthodox interpretation of the nature of black male socialization in matrifocal families" [p 4]) that black youth in this setting are socialized into ineffectual, deviant male roles is systematically questioned by the articles which follow. Ulf Hannerz points out in "Growing Up Male" that fatherless households are rarely devoid of male role models; by the same token mothers frequently instruct boys in so-called manly behavior. Charles Kiel develops the cultural aspects of the ghetto male role more thoroughly in his article, "The Expressive Black Male Role: The Bluesman." Robert Coles's examination, "Black Fathers," reveals the often ignored majority of poor black fathers who struggle against the severest odds in their efforts to care for their families. Quite distinct

from the orthodox view, these men are shown to be strong, proud, loving—and loved—figures in their families. Finally, Seymour Parker and Robert Kleiner's study, "Social and Psychological Dimensions of the Family Role Performance of the Negro Male," identifies societal discrimination in employment, housing, and general opportunities rather than subcultural norms as the major determinants of black male family role performance.

William Turner's "Myths and Stereotypes: The African Man in America" opens the second section with an in-depth analysis of the historical bases of myths and stereotypes about black men. In the process he offers a revealing critique of scientific modeling as it has been applied to the study of black males. Robert Staples systematically dissects and refutes two of the many stereotypes discussed by Turner in his articles entitled "The Myth of the Impotent Black Male" and "The Myth of the Black Matriarchy." It is Staples's conclusion that although both concepts are common to the literature, they lack legitimate bases in either theory or fact. Similarly, Harry Edwards lays to rest notions of black biophysical superiority—or inferiority, depending on one's perspective—in his article, "The Sources of the Black Athlete's Superiority." He finds black "superiority" in this arena to stem from the denial of equal opportunities to blacks in other fields of endeavor. Last, Doris Wilkinson's "The Stigmatization Process: The Politicization of the Black Male's Identity" offers a thoughtful consideration of color stigmatization processes and their detrimental consequences for the status of black males.

Beth Day's article, "The Hidden Fear," introduces the touchy topic of interracial mating by probing the hidden fears associated with black male–white female mating in our society. James Blackwell's "Social and Legal Dimensions of Interracial Liaisons" examines institutional barriers to increased interracial mating and predicts that as these barriers diminish rates of black-white dating and marriage will rise. Beth Downs supports such a view in her "Black/White Dating" when she links the dramatic growth in transracial dating and marriage to political liberalization of the society during the sixties. Selections by Calvin Herton ("The Negro Male") and Doris Wilkinson ("Expectations and Salience in White Female–African Male Self-Other Definitions") serve, however, as reminders of the continuing negative stereotypes and degrading stigmas associated with such unions. The taboo status of white females, mythical portrayals of black male supersexuality, and the white female's cultural right to establish the rules of black male–white female mating games are all examined in this context.

The Black Male in America concludes with an examination of contemporary and projected future statuses of black males in American society. Ronald Taylor's insightful opening paper assesses the causes and consequences of the continuing concentration of black workers in the lower reaches of the occupational hierarchy—a condition he fully expects will worsen over the coming years. He also expresses related concern over the "deepening schism" in the black community as middle- and upper-class salaries more rapidly exceed those of the much larger black underclass.

Nathan Caplan, in "The New Ghetto Man: A Review of Recent Empirical Studies," speaks to some political consequences of continued deterioration in the overall economic status of blacks and concludes that violent disruption of urban centers is a constant possibility. In keeping with the book's iconoclastic bent, Edwin Harwood and Claire Hodge's "Jobs and the Negro Family: A Reappraisal" convincingly challenges the myth of black female economic superiority over black males. Black females are shown to be the most exploited of all the economically deprived. Charles Willie's "The Black Family and Social Class" and David Schulz's "Variations in the Father Role in Complete Families in the Negro Lower Class" comment on the often neglected diversity of black families. These two papers show that family structural and interactional patterns vary across as well as within classes.

As the summary above suggests, this book is worthy of careful attention. If I were to criticize it, it would be for whetting my appetite and then leaving me unsatiated. After reading the book a host of questions were provoked in my mind concerning the status and roles of black men in this country's criminal justice, religious, military, and political systems. What also of the relative value of competing theoretical perspectives for understanding black males? Or, which alternative research designs should we strive to apply in the future? Are there lessons of importance to be shared with legislators and other policymakers whose decisions impact upon black males? The questions stimulated but left unanswered by this volume are endless. But then, this is perhaps as it should be. Wilkinson and Taylor have assisted in organizing our thinking about topics of relevance to the study of black males and have provided heuristic motivation for future research in the area. For this significant contribution they are, I think, to be congratulated. It now remains for those of us with interests in this area to respond to this impetus by more extensively researching black male statuses and roles in our society.

Turkish Workers in Europe, 1960–1975: A Socio-Economic Reappraisal. By Nermin Abadan-Unat and contributors. Leiden: E. J. Brill, 1976. Pp. viii+424. Fl 160.

Y. Michal Bodemann
University of Toronto

Over the past decade or so, the principal focus of studies in labor migration has been on three major geographical areas: (1) migration from Mexico to the Southwestern United States, which as farm labor has a strong seasonal character; (2) migrant labor in Southern and Eastern Africa, which is marked by racially defined wage differentials and "labor aristocracies"; and (3) current West European labor migration. In his paper "The Func-

tions and Reproduction of Migrant Labor" (*AJS* 81 [March]: 1050–87), Michael Burawoy has presented an interesting comparative analysis of the African and American cases, and S. Castles and G. Kosack (*Immigrant Workers and Class Structure in Western Europe* [London: Oxford University Press, 1973]) have dealt comprehensively with the situation in Western Europe.

One of the distinctive characteristics of the European case is the national divisions within the foreign work force. In West Germany in particular it had originally been a policy to recruit roughly equal shares of foreign workers from the various nationalities—Italy as an EEC member excepted—and to disperse these groups evenly throughout the country. Most work places and most company-owned housing reflect this dispersal. Commonly, two or three foreign nationalities are heavily represented, with sprinkles of the remaining groups. However, this balanced recruitment has been upset by the massive influx of Turkish and Yugoslav workers beginning in the late 1960s. Clearly, an attempt was made not only to separate native workers from foreigners, but also to stratify the foreign labor force internally, with the Turkish workers at the bottom of the wage scale and work hierarchy, and to pit one group against the other. It is interesting to note that this policy, aimed at preventing working-class solidarity, has not worked very well: foreign workers have supported most labor disputes and have often themselves initiated strikes, as in the case of the so-called *Türkenstreik* at Ford in Cologne in 1973, which was brutally suppressed by the police with help from their friends, the German labor unions. Unfortunately, not a single contribution in the present volume addresses this entire complex of questions.

A second important characteristic of the West European case is the type of its reproduction of foreign labor. If we use Burawoy's distinction of renewal and maintenance as the basic elements of reproduction, we see that there is a significant degree of variation among the different countries. Switzerland, at one extreme, separates processes of maintenance and renewal of labor most sharply: with an extraordinarily high ratio of foreigners in its labor force (roughly 30%, five times that of the other major European countries), it maintains a high degree of labor rotation, makes it virtually impossible for foreign workers to bring their families, and discourages vocational training and other forms of education. In this sense, it reminds one most clearly of the case of the South African mine workers. In complete contrast stand Belgium and Sweden, which tend to favor absorption of the immigrant into the new environment. By and large, however, Germany and France, with the bulk of the total foreign labor force, appear to be most representative of the European case: part of the foreign labor force is rotated, the other part shows a high degree of permanency, and a—very limited—attempt is being made to provide schooling or to offer vocational training. It is important to see therefore that, unlike in "pure" cases of migrant labor, in Western Europe some functions of renewal of labor are being shifted from the sending country to the receiving country.

Originally the influx of foreign labor at large helped to depress the wage level; today this task is assigned to the rotation segment (and the mass of illegal immigrants) in most EEC countries. It appears that this function of providing a rotating labor force was given primarily to Yugoslavia and Turkey. Several contributions by both German (Otto Neuloh) and Turkish (Turhan Oğuzkan) authors in the present volume directly advocate rotation even today, while little attention—except, in part, the paper by Ger Mik and Nia Verkoren-Hamelaar on foreign worker segregation in two Dutch cities—is given to the realities of relatively stable ghetto-type settlement, particularly of Turkish workers, in the receiving countries.

The case of European labor migration is further peculiar in that it forces an analysis of the internal character and the functions of the sending countries—a point often neglected in other regions of migration. This is perhaps the main focus of *Turkish Workers in Europe, 1960–1975*. Before the onset of its labor export, Turkey, like most other sending countries, underwent a sharp internal polarization between urban centers and the rural periphery. Unfortunately, Abadan-Unat's collection does not address itself to the causes of this polarization—for example, the agricultural policies of successive Turkish governments. Most of the contributors are content with the explanation that overpopulation as "push" factor just happened to meet the "pull" factors in the core countries of the EEC. Only Mübeccel Kiray, in a paper on the migrant worker's family structure, points briefly to the increasing indebtedness of the peasant as a cause of emigration.

While the conditions leading to emigration are not dealt with in a serious manner, several contributors point to the more visible concomitants of Turkey's internal polarization, especially the rural-urban migration within the country itself. This movement began roughly at the end of World War II and prepared the way for the international migration. Many of the migrants first moved from the rural to the urban areas and then moved on to central Europe. Furthermore, as Ruşen Keleş points out, roughly 50% of all foreign workers who left rural regions and went directly to Europe do not return to their hometowns later but settle in the larger cities. This process may be particularly pronounced in Turkey; it can also be observed in other sending countries such as Italy, Spain, and Greece.

Whereas the role of the sending country is to mobilize and export cheap labor to be consumed by EEC capitalism, it is the role of the migrant worker to open up markets, particularly by means of remittances, for the import of the commodities which they themselves help produce. Although foreign workers are underpaid compared with native workers, their high amount of overtime work, tax-free moonlighting, and their frugality, all of which can be sustained for only a limited time period, allow them to accumulate savings. In contrast to what appears to occur in the case of the South African mine workers, however, these savings are only partly used for immediate renewal functions—family subsistence—at home, and mostly for the purchase of consumer goods or unproductive investments back home,

as Abadan-Unat convincingly points out. While remittances temporarily improve the balance of payments of the sending country, they in fact help to increase the dependency of the sending country on West European capital. Some of this is dealt with reasonably satisfactorily in *Turkish Workers in Europe,* most notably in Günter Schiller's intelligent contribution.

In sum, this is a rather disappointing survey of the Turkish foreign labor question. The bulk of the contributions, often repetitive in substance, is replete with social policy clichés and well-intentioned banality ("The long-term goal must be to put an end to economically forced manopower [*sic*] movement" [p. 188]). Going through the volume, I have often wondered whether most of the contributors have ever ventured outside academic walls and actually talked to a migrant worker. Particularly appalling is the sloppy paper by O. Neuloh, who finds that people in underdeveloped countries do not like to engage in "handicraft and heavy industrial work" (pp. 53–54) and prefer office jobs, or that some aspects of Turkish backwardness must be attributed to its "national religious and biological mentality" (p. 72). Only M. Kiray presents a truly insightful study which points to the radical readjustment of migrants' kinship structures and to the drastic change in the role and the social status of women in migrant workers' families.

Finally, a word about the miserable job of editing this high-priced volume. I believe I have never seen anything in print so riddled with typographical and other errors. To mention just one example: "The labour market in industrial society will be determinated [*sic*] by the percentage of labour force [*sic*] depending in renumeration [*sic*] in wages and salaries" (p. 51). I hope this renowned international publisher is sufficiently embarrassed about his shoddy job.

Strangers in Their Midst: Small-Town Jews and Their Neighbors. By Peter I. Rose, with the assistance of Liv Olson Pertzoff. Merrick, N.Y.: Richwood Publishing Co., 1977. Pp. xiii+224. $10.95.

Samuel C. Heilman
Queens College, City University of New York

As Marshall Sklare once noted, "Jews, more than any other group, have been among America's most enthusiastic city-dwellers" ("Jews, Ethnics, and the American City," in *Commentary* 53 [1972]: 70). Not only have they themselves adapted to urban life, they have also evolved a major role in supplying others with many of its amenities. Nevertheless, for most Jews such an urban existence was never a repetition of their European experience. There they lived in the small, rural village—the *shtetl.* Yet although many came from such villages, they seldom idealized that rural experience,

for they realized that in their *shtetl* existence they had occupied the bottom rung in the social order. The American city, on the other hand, symbolic of a new world where achievement rather than ascription was the determinant of social status and fortune, offered a new start. However difficult that start might be, most American Jews elected to be urban rather than rural in outlook and location.

In the late fifties, Peter Rose, then a graduate student in sociology at Cornell, decided to examine and interview those few Jews who, unlike 98% of their brethren in America, had chosen to live in small towns. He selected to study particularly those Jews living in the rural villages of upstate New York, the area most accessible to him. Twenty years later, one of Rose's students returned to the area, and to many of the families originally studied, for a brief follow-up. What emerges from this joint effort is a portrait of an American Jewish experience which was "swift as a shadow and short as any dream." The world of the small-town Jews of upstate New York is no longer; their children have gone to the cities.

At the time that Rose first talked to these rural Jews, their world was already in its decline, indeed seemed never to have really been on the rise. Unlike the rural world of the European *shtetl* which Mark Zborowski and Elizabeth Herzog poetically described as one where "life is with people," the world of the American village was for the Jew a "life without people." Again and again one reads that these people viewed themselves as "an isolated minority" and as "strangers." The rural Jew, living in a town with at most nine other Jewish families, missed the company of *his* people, and worried that his children might intermarry with Gentiles (although in the follow-up it turns out that few did). However successful economically and professionally the small-town Jew became, he saw himself relegated to the ascribed role of "ambassador," one who provides an acceptable model of his people for others. As Louis Wirth once put it, "getting out into the Christian world made the Jew self-conscious" (*The Ghetto* [1928; Chicago: University of Chicago Press, 1956], p. 123).

Yet instead of becoming "exemplary," conceived by their Gentile neighbors as representative of Jews in general, these Jews were viewed by those among whom they lived as "exempted"—different and special. Only occasionally did Rose discover evidence that Gentile attitudes toward the Jews were improved as a result of an intimate and positive experience with a particular Jew. Over 80% of the Christian opinion leaders whom Rose polled in his small towns continued to think of Jews in terms of the negative stereotypes of shrewdness, aggressiveness, clannishness, and the like.

Perhaps the particular questions that Rose asked emphasize the point, but the small-town Jew that emerges from these pages is one who constantly expresses—either directly or indirectly—a desire for legitimation by the Christian world. This Jew feels the urge to try to define his position vis-à-vis his non-Jewish environment. He explains his ritual and holy day observance, such as it is, in Christian terms. Passover becomes, for example, "a sort of Jewish Thanksgiving." He practices only those rituals which can

be made somewhat sensible in the American Christian context. He describes himself in terms which belie anti-Semitic, Christian stereotypes; Jews are "not underhanded or shrewd." In short, whether happy or not in the small town he inhabits, he acts as if he were there by the consent and sufferance of the dominant Gentile culture. Any ethnic pride he may have is muted in its presentation.

The lesson of this life was not lost upon his children. Most have left, not really to the surprise of their parents. Nearly half those polled by Rose expressed a desire to have their children leave the small town. That is, they realized that however adjusted they themselves may have become to rural American life, the American small town is not a place where the Jewish people may feel at home.

What had drawn them there had been economic needs. In the small town a peddler could more swiftly transform himself into a merchant, and many did. In the small town a refugee physician could more easily practice medicine, and many did. But the price for this economic and professional success was often an overwhelming sense of alienation. Even those who expressed a satisfaction with rural life felt the constant threat of being swallowed up by that Christian world that had allowed them entry. "Of all the problems the isolated Jew must face, none is more pressing than that of preventing his children from becoming *goyim,*" Rose tells us (p 151).

If one is to judge by Rose's sample, these Jews were at least successful in avoiding the apostasy of their progeny. Their children seem to have followed the pattern predicted by the historian Hansen: "what the son wishes to forget the grandson wishes to remember." They are aggressive and contra-acculturative in their Jewishness. Not only have many left the small town, they have begun in significant numbers to openly practice their Judaism and to seek publicly to identify themselves distinctively as Jews.

But, one wonders, is this the true story of small-town Jews? Rose's method of collecting his sample, snowballing (asking rural Jews to give names of other rural Jews they know), seems to assure that he would speak to those who remained bonded to the Jewish community, however tenuously. What of those Jews who left completely, who moved to the West and the South? What of the Goldwaters, and others like them, who so integrated themselves with rural America that they ceased to hold any identity as Jews? These people, even more than Jews in rural New York state, were the small-town Jews of America. These seem to forever have been lost.

Nevertheless, if Rose and Pertzoff have not told the entire story of the American small-town Jews in *Strangers in Their Midst,* they have given us a unique glimpse of one American Jewish character. Perhaps most interestingly, this description of a deviant Jewish life form is further evidence that in America, as in Europe, in the final analysis, "a Jew," as the philosopher Nathan Rotenstreich once put it, "is a Jew only when he is with other Jews" ("Emancipation and Its Aftermath," in *The Future of the Jewish Community in America,* ed. D Sidorsky [New York: Basic, 1973], p. 56).

Community and Polity: The Organizational Dynamics of American Jewry. By Daniel J. Elazar. Philadelphia: Jewish Publication Society of America, 1976. Pp. xv+421. $12.00.

Marshall Sklare
Brandeis University

Daniel Elazar belongs to that small minority of Jewish social scientists who have very strong Jewish commitments. He is distinguished by the fact that he has both an intimate knowledge of contemporary Jewish affairs and an acquaintanceship with classical Jewish sources. Elazar is a rare individual: a scholar and an activist, a general theorist as well as a political scientist passionately devoted to the furtherance of contemporary Jewish studies, an academic administrator as well as a scholar who is both a teacher and a researcher.

All of Elazar's diverse interests and qualities are evident in *Community and Polity.* The result is a book which is unique and important. The true significance of *Community and Polity* is apparent when it is compared with the only other book on the structure of the American Jewish community: Maurice J. Karpf's *Jewish Community Organization in the United States* (New York: Bloch). As Karpf's book was published in 1938, it contains no information on the events of the past four crucial decades. In addition, Karpf's book is limited by being written from a social work perspective—in fact, the impetus for the book was the Second International Conference on Jewish Social Work, which was held in London in the summer of 1936.

Elazar covers some of the same ground traversed by Karpf, but he does so in an entirely different way. Instead of systematically treating which organization delivers what social services, he analyzes the principles of organization on which the American Jewish community is based. Painting with broad strokes, Elazar excites interest by moving away from the social work approach which has characterized writing on the Jewish community. His book is novel because he writes about the Jewish community in the same way that political scientists write about community in general. With the exception of the works of Charles Liebman, this type of analysis has been sorely lacking in writing on the Jewish community.

Elazar's book covers a host of topics. He analyzes the degree of pluralism in American society and studies the volunteerism on which the American Jewish community is based. He analyzes the American synagogue in its diverse forms. He provides a great deal of information on the federations—the fund-raising agencies found in all cities with a Jewish population of any substantial size. The federations subsidize local services and act as a central address for the Jewish community. A substantial part of the receipts from their annual campaigns is allocated to national and international Jewish organizations and to the needs of Israel.

Elazar attempts to delineate the process of decision making in the federations and in the American Jewish community in general. He analyzes

what he calls the "trusteeship of givers and doers" (p. 336). Furthermore, he describes the uniqueness of the various regions of the nation and the special characteristics of Jewish communal life in the cities located in these regions. He also proposes a scheme for analyzing the relationship of a given individual to the community. Discerning seven levels of participation in communal life, he provides an estimate of the proportion of Jews on each level. Thus, Elazar has written a volume which will serve the purpose of introducing social scientists to the phenomenon of the organized Jewish community. *Community and Polity* will also be a book which no student of contemporary Jewish studies can afford to ignore. Finally, it will undoubtedly serve to raise the level of writing and discussion about Jewish communal affairs.

The defects of *Community and Polity* are two. The first is a combination of carelessness, incompleteness, and sloppiness. For example, although Elazar devotes several pages to a discussion of the Jewish community center as an institution in American Jewish life, his index does not contain the entries "community center, Jewish" or "Jewish community center." Elazar's footnotes are replete with errors. In a single sentence on page 406 there is a typographical error in the last name of one Jewish educator (Judah Pilch) and in the first name of another (Alvin Schiff). Appendix D, entitled "Jewish Community Studies: A Selected Bibliography," obviously requires a period of uninterrupted work in a library with a large Jewish collection to bring it up to standard. For example, in his section on Boston, Elazar fails to list the standard book in the field: *A Chronicle of Boston Jewry* (Boston, 1963) by Albert Ehrenfried. The section on Chicago has the same defect—Elazar fails to list the standard book: *A History of the Jews of Chicago* (Chicago: Jewish Historical Society of Illinois, 1924), edited by H. L. Meites. Elazar also takes delight in breaching confidentiality, his practice is to identify the true name of communities for which authors have constructed pseudonyms.

The second defect of *Community and Polity* is ideological and thus cannot be remedied so easily. Elazar's central view is that two institutional forms have emerged as predominant in the American Jewish community: the synagogue and the federation. Elazar is well acquainted with the American synagogue and is sympathetic to its purposes. He is, however, troubled by its localistic orientation and by its lack of articulation with the American Jewish polity as a whole. The end of the matter is that he views the federations as the center of the Jewish community. His picture of the federations is almost entirely positive. It is, in fact, out of keeping with what might be expected of a treatment of this institutional form by an academician.

Although the federations have been quite successful in raising large amounts of funds, it is apparent that they have avoided addressing basic problems of the Jewish community since most of these problems are not philanthropic in character. Federations raise large sums for Israel. But Israel's problems are political, and the crux of American Jewish responsi-

bility is the task of mobilizing political support for Israel. The problem of Soviet Jewry is also a political one. Although the funds raised by federations have been helpful to Soviet Jews who have settled in the United States and in Israel, the federations have been peripheral to the effort to fight restrictions on the right of emigration from the Soviet Union and to improve conditions for those Jews who wish to remain there. On the local scene the most obvious problem is that of decaying Jewish neighborhoods. Since this problem too is basically political, federations have either been passive or impotent. Whatever strategies have been utilized to stabilize Jewish neighborhoods have come from local residents and from neighborhood-based Jewish organizations rather than from the federations.

Elazar is surely aware of all of these problems, but they do not seem to give him pause. He does not address himself to whether the federations take on only the problems which they feel are manageable and leave to others the really difficult questions of Jewish life. Elazar also does not discuss whether the solution to whatever issues federations do take on is invariably the traditional Jewish response of *tzedakah,* or what may be loosely translated as philanthropy.

There can be no objection to an academic becoming an activist. It is refreshing to see the example of an Elazar who does not feel the alienation which academicians often experience when they work very closely with the "trusteeship of givers and doers." On the other hand, *Community and Polity*—with all its admitted excellences—leaves us with the question of whether the social scientist should not keep a certain distance between himself and the politics of the community lest he be attracted to the conventional wisdom.

World of Our Fathers: The Journey of the Eastern European Jews to America and the Life They Found and Made. By Irving Howe, with the assistance of Kenneth Libo. New York: Simon & Schuster, 1977. Pp. xx+714. $6.95 (paper).

Jack Nusan Porter
University of Lowell

This book (at times it seems better to call it an encyclopedia) is a Jewish *Roots,* just as *Roots* is a black *World of Our Fathers.* Both Alex Haley and Irving Howe have achieved what is probably the goal of nearly every sociologist—to write a serious sociohistorical narrative that becomes a nationwide best seller. Now, if only a sociologist could accomplish this, it would give our shopworn profession a shot in the arm!

Both Haley and Howe have combined primary scholarly material (early newspapers, shipping records, immigration and governmental reports, first-person memoirs) with a compelling novelistic writing style to produce a

book that is both a joy to read and a contribution to historical scholarship. Here I will only touch on a few sociological highlights.

Howe writes with supreme confidence and with a total grasp of his subject matter. He combines a knowledge of Yiddish literature and literary criticism with a knowledge of Jewish contributions to socialism and trade unionism. To be at home in both literature and politics is a rare and delightful combination. The book flows effortlessly from the persecution of *shtetl* Jewry in Eastern Europe, through their passage to America, their settlement on the Lower East Side of New York (which at the turn of the century had a population density that equaled the worst sections of Bombay), their struggle for success and dignity in the *goldineh medineh* ("golden land of opportunity"), their contribution to American culture, to the unease of the children and grandchildren, who still struggle to maintain their identity as Jews. It is the story not only of the Jews but of every immigrant group—and every immigrant group should find a Howe or a Haley to record its journey to America.

Howe's strongest chapters are those dealing with the Yiddish press, the socialist/trade unionist movements, and the role of Jewish scholars, intellectuals, and entertainers in American life. As a sociologist, however, I was irritated by his obvious blind spots regarding the role that sociologists (especially the Chicago School) played in such areas as social reform, immigration policies, and the study of immigrant pathology.

To be sure, he does mention some people briefly, for example, social historians such as Jacob Lestchinsky, Elias Tcherikower, Mark Wischnitzer, and Jacob Shatzky; essayists and philosophers such as Hayim Greenberg, Morris Raphael Cohen, and Hayim Zhitlovsky; Holocaust writers such as Isaiah Trunk, Philip Friedman, and Jacob Robinson; and Labor Zionist political philosopher Nachman Syrkin and the sociologist C. Bezalel Sherman. Most sociologists have probably never even heard of these people. Furthermore, many sociologists may also not know (as I myself did not) the answers to several questions that Howe brings up: What side did sociologists take on the immigration controversies? Did they favor "open admissions" or restrictive quotas? What were the prevailing political attitudes of sociologists toward reform, pathology, etc.? They studied immigrants, and quite well, but I am curious about their personal attitudes on the political issues of the day. Howe's book tantalizingly opens up such avenues of inquiry.

However, he makes no mention of the important contribution of the Chicago School to immigrant studies. Prominent members of the school were intrigued by the Jewish immigrant: W. I. Thomas (who is not mentioned) did a study of the "Bintel Brief" ("Bundle of Letters") column found to this day in the *Jewish Daily Forward*, the largest Yiddish paper in the world, and he learned Yiddish to make the study. Ernest Burgess, Ellsworth Faris, and Robert E. Park encouraged and guided Louis Wirth in writing the now classic *The Ghetto*, which is never mentioned by Howe. There is work for sociologists to do in this area where Howe barely treads.

He is very good in describing the "social dislocation and pathology" (p. 96) of the Jewish quarter on the East Side. He practically entices the reader to go back in history to verify or refute his generalizations. For example:

> There was crime, there was wife desertion, and there were juvenile delinquency, gangsterism, and prostitution during the eighties and nineties, as well as during the early decades of the twentieth century—probably more than the records show or memoirists tell. How could there not be?
>
> Precise information on these matters is hard to come by, and the reasons are obvious. Communities struggling for survival seldom rush to announce their failures. The craze for sociological investigation that would overtake America in the twentieth century was not yet very strong. And over the centuries the Jews had developed a cultural style encouraging prudishness and self-censorship: there were things everyone knew, had no choice but to know, yet only rarely was it deemed proper to speak or write about them. Life was hard enough without indulging in luxuries of revelation. [P. 96]

> The single greatest journalistic quality of the [*Jewish Daily*] *Forward* was the sustained curiosity it brought to the life of its own people. In a later, more portentous age, this would have been called "the sociological imagination." . Were East Siders finding new occupations ranging from real estate to gangsterdom? Were lonely immigrant girls succumbing to the lure of suicide? Were *yentes* (busybodies) moving to West End Avenue and becoming "fancy ladies"? . . . Nothing seemed too mundane for the *Forward* staff, and that may be one reason high-minded Yiddish critics often treated the paper with contempt. [P. 531]

It is interesting that Howe's strongest areas have come under the most attack from some Yiddish quarters, from people with long memories. Paul Novick, the editor of the *Morning Freiheit*, the oldest Jewish-Communist paper in the world, has recently written a booklet entitled "The Distorted 'World of our Fathers' " (*Morning Freiheit*, New York, 1976), wherein he condemns Howe in part for his errors regarding some of the early trade union strikes but moreover because he does not berate Abraham Cahan (the founding editor of the *Forward*) enough for vulgarizing Yiddish by introducing too many Englishisms into the language and for ridiculing "highbrow" Yiddish literature (e.g., Sholem Aleichem). The old struggles, unknown or forgotten by the children, are still remembered by the old-timers.

But, even if true, these are really minor criticisms for a book that tells so well the story of *Yiddishkeit* and *Yiddin* in the New World. I recommend it to every social scientist, and, now that it is out in paperback, it is affordable for students as an auxiliary textbook.

Ruling Class, Ruling Culture: Studies of Conflict, Power and Hegemony in Australian Life. By R. W. Connell. Cambridge: Cambridge University Press, 1977. Pp. xii+250. $19.95.

Ivan Szelenyi
Flinders University of South Australia

R. W. Connell is certainly one of the most interesting and stimulating personalities in Australian sociology. His first significant work was a study (à la Piaget) of the early political socialization of children. By the mid-seventies he apparently converted to Marxism and discovered the Australian ruling class for the Australian academia. His short, well-written essay (theoretically not much beyond an elementary Marxism-Leninism seminar), "Yes, Virginia, There Is a Ruling Class" (with T. Irving, in *Australian Politics*, ed. H. Meyer and H. Nelson [Melbourne: Cheshire, 1973]), is still compulsory introductory reading for all would-be radical sociology students on Australian campuses, and with its publication Connell at once became the *enfant terrible* of the sociology establishment. Certainly in the seventies it would have been difficult to shock anybody in the United Kingdom or in the United States with the rediscovery of class analysis, but Australian sociology lagged behind the mainstream of research. With the exception of Sol Encel, whose *Equality and Authority* (Melbourne: Cheshire, 1970) is still the only comprehensive analysis of Australian social structure, few sociologists in Australia showed much interest in macrostructural questions. Thus Connell moved into a theoretical vacuum and, surrounded by admirers and enemies, attempted as a one-man task the almost impossible—to rewrite Australian sociology from the point of view of class analysis. Readers of *Ruling Class, Ruling Culture* should bear in mind the pioneering character of this work.

Ruling Class, Ruling Culture is a collection of essays written between 1972 and 1977. These essays are partially work-in-progress papers in which Connell attempts to understand the meaning of class and the differences between class and stratificational analysis, and partially writings for a more general audience on such wide-ranging topics as the short-lived rule of the Labor Party, the nature of the mass media, a 15-page sociology of Australian big companies, and a similarly condensed overview of the problems of class consciousness in childhood. These later essays were written with the obvious purpose of demonstrating to the wider public that the new sociological approach is justified and helps to explain the most varied phenomena. Professional sociologists, therefore, should not expect to learn much about the actual "conflicts, power and hegemony in Australian life" (although this is what the subtitle of the book promises). Rather, I think we should look upon this collection of essays as preparations for a more comprehensive work. These papers should help Connell to write and his readers to appreciate his still unfinished book on the class history of Australia.

In the first two essays, which I called work-in-progress papers above, Connell offers a criticism of the stratification approach to social structure. There is nothing spectacular or original in this critique (it is not very easy to be spectacularly original in such an area so much worked upon), but Connell's distinction between class structure and stratification is clear and precise and does not lack a certain degree of sophistication: he identifies class theory as a "generative" one, stratificational theory as a "categorical" one. One only wonders why an author, who is thus aware that we are dealing with analyses on different levels of abstration and with different purposes, engages himself in a full-scale attack against stratification research. Connell makes the same mistake as his stratificationalist opponents who wanted to replace class with stratification. Connell accepts this false either-or dilemma. Instead of explaining to us what a class analyst can do with data on social stratification, he rejects stratificational analysis altogether.

Connell's own understanding of class has undergone a significant change since "Yes, Virginia, There Is a Ruling Class." From the two first chapters one gets the impression that he still operates within the classic Marxist paradigm, but from the more empirical essays that follow it is clear that the author moves away from a class definition based on ownership and increasingly treats power as a class-generating factor; he moves from Marx toward Dahrendorf. The reader is not warned of this significant shift in theoretical orientation, and one has to be quite familiar with different class theories and their cobwebs to realize what the author is doing here. Connell tries to avoid a narrow definition of the ruling class and thus proposes "an extension of the initial definition of ownership and the concept of profit" (p. 49). However, he finally decides not to attempt this extension on the theoretical level but to look empirically at the managerial elite, at the leadership of the Liberal and Country party, and at the top civil servants. He classifies them (partly by reverting to the logic of stratificational analysis) into the ruling class, basically on the grounds that they exercise power in the interest of capital. In an implicit way (for he does not attempt to interpret the notion of the working class) we are confronted with a class dichotomy between the rulers and the ruled, the powerful and the powerless, which seems quite a few steps away from Marxist class analysis.

One hardly can question the political relevance of this theoretical turn. In addition, Connell does not hesitate to mix a bit of psychoanalysis into his class theory if it is necessary in order to strengthen his argument. This is most obvious in the last essay, "Pattern of Hegemony," which may also be the most original in the book. Here Connell applies with a great deal of imagination Gramsci's concept of hegemony, although he probably overstates the deterministic overtones of this notion.

Connell's overtures to theoretical eclecticism, however, worry me. As far as I can judge, in this move he shares the growing pains of many left-wing social scientists in Australia for whom the theory is nothing, the practice is

rything. We should be aware that the subordination of theory to the ds of political practice is the end of theory.

Ruling Class, Ruling Culture is not a great book, but it is a stimulating and I can warmly recommend it even for use in undergraduate teaching e. We are looking forward to Connell's next book on Australian class ory, which I expect will show more theoretical rigor and be based on e solid empirical evidence.

RURAL HEALTH CENTER SERIES

Health Services Research Center, University of North Carolina, Chapel Hill, and the Office of Rural Health Services, North Carolina Department of Human Resources

This pioneering six-volume series is the only comprehensive guide available for the establishment of rural health centers for primary care. It provides New Health Practitioners (as well as community leaders and physicians) with detailed informatio for organizing and operating these facilities. The authors combine their own knowledge and experience with a survey of more than 100 operating rural health centers to insure that both new and existing centers provide maximum benefit to the community. Extensively illustrated, this series represents the first unified approach to giving America's rural communities the improved health care that they have needed.

Vol. I - RURAL HEALTH CENTERS IN THE UNITED STATES
James D. Bernstein, Frederick P. Hege, Christopher C. Farran
April 1979 ca. 96 Pages
33 Photographs 22 Illustrations
Cloth $10.00 Paper $5.95

Vol. II - PLANNING AND MANAGING RURAL HEALTH CENTERS
Edward F. Brooks & Torlen L. Wade
April 1979 ca. 128 Pages
25 Photographs 43 Illustrations
Cloth $15.00 Paper $7.95

Vol. III - FACILITY PLANNING, DESIGN AND CONSTRUCTION OF RURAL HEALTH CENTERS
Terry W. Alford
April 1979 ca. 256 Pages
155 Photographs 63 Illustrations
Cloth $17.50 Paper $9.95

Vol. IV - A LEGAL GUIDE FOR RURAL HEALTH PROGRAMS
David G. Warren
April 1979 ca. 192 Pag
40 Photographs 25 Illustratio
Cloth $15.00 Paper $5.9

Vol. V - CLINICAL ROLES IN RURA HEALTH CENTERS
John W. Denham & C. Glenn Pickard
April 1979 ca. 96 Pag
25 Photographs 6 Illustratio
Cloth $10.00 Paper $5.9

Vol. VI - MEDICAL RECORD AN INDEX SYSTEMS FOR COMMUNIT PRACTICE
Robert J. Sullivan
April 1979 ca. 96 Pag
43 Photographs 10 Illustratio
Cloth $12.50 Paper $6.9

EVALUATION IN MEDICAL EDUCATION

Past, Present, Future

Edited by Thomas Samph, Ph.D. and Bryce Templeton, M.D., National Board o Medical Examiners and American Board of Family Practice

Possible future roles of organizations and institutions connected with the profes sional responsibilities of physicians with regard to evaluation of medical education.

ca. $20.0(

THE IMPORTANCE OF MENTAL HEALTH SERVICES TO GENERAL HEALTH CARE

Jeffrey L. Houpt, Carole S. Orleans, Linda K. George, and H. Keith H. Brodie, Duk University Medical Center

Examines the relationships between mental health services and general health care i three areas: psychological disorders in general medical practice; behavioral risk fac tors associated with major illnesses and efficacy of available intervention strategies the interface between biological, psychological, and lifestyle or behavioral factors ir illness onset and treatment.

In production ca. $20.0(

mber.